CHOICES AND PREFERENCES AT THE INDIVIDUAL, HOUSEHOLD, AND COMMUNITY LEVELS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Xin Gao

May 2019

CHOICES AND PREFERENCES AT THE INDIVIDUAL, HOUSEHOLD, AND COMMUNITY LEVELS

Xin Gao, Ph.D.

Cornell University 2019

This dissertation consists of three essays that examine choices and preferences at the individual, household and community levels. The first chapter, "Labor Market Outcomes with Heterogeneous Preferences and Search Frictions: The Case of Chinese Migrant and Urban Workers", examines decision-making at the individual level. Neither Rosen's classical compensating differential model nor newly developed search models could explain the particular pattern of wage and job characteristics distributions in China, where migrant and urban workers coexist and dominate different sectors. In the theoretical part of this paper, I expand the model by Lang and Majumdar (2003) to show that wages need not be compensating when preferences are heterogeneous, and that the group more averse to undesirable working conditions need not earn less when reservation utilities differ and/or when employers practice taste-based discrimination. In the empirical part, I substantiate the assumption in my theoretical modeling that urban workers are more averse to undesirable working conditions using a discrete choice experiment, where 225 workers in China made hypothetical choices between jobs characterized by different wage levels and working conditions. I backed out preference parameters and willingness to accept measures for job attributes. I find that consistent with my assumption, urban workers need to be compensated more to accept outdoor jobs and jobs in second line cities. I also find that migrant workers have more dispersed preferences that vary with personal characteristics such as gender and

education.

The second chapter,"Where Did the Money and Time go? De-mystifying the Negative impact of Remittances on Human Capital Investment in the Kyrgyz Republic", examines how households allocation of financial and labor resources affects human capital formation in the Kyrgyz Republic. International migration and remittances from overseas may encourage human capital investment and improve educational outcomes in developing countries. Empirical studies, however, have shown mixed evidence. In our study, we focus on the case of Kyrgyz Republic, one of the largest remittances-receiving countries in Asia. I used a 5-year panel dataset that tracks the same 3,000 households and 8,000 individuals in the country to examine the impact of remittances on both household educational expenditure and attendance rate of school age children. I used instrumental variables and fixed-effects regressions to correct for potential selection bias. I find that remittances have a negative impact on human capital formation – namely both educational expenditure and school attendance rate are lower for households that receive a higher amount of remittances. To explore the possible channels of the negative effects, I further regressed itemized household expenditures and the time use pattern of school-age children on remittances. I find that the negative effects can at least be partly attributed to increased expenditure on durable goods and increased hours of child labor on farm work as a compensation for adult labor insufficiency induced by out-migration. My finding calls for the monitoring of farm labor hours of school-age children. Moreover, implementation and scaling up of financial literacy programs that help parents balance short-term expenditures (durable goods) and long-term investments (education, health) can be beneficial. In addition, targeted investment to improve the quality of education services in the country may help increase perceived return to schooling and may therefore improve human capital investment.

The final chapter,“Kinship, Social Preferences and Voting in Rural China: A Lab-in-the-Field Experiment”, goes beyond individuals and households to examine communities and how social preferences and social network can affect collective decision-making. Economists have come to understand that human choices are not only driven by self-interest but also “social preferences” – a person’s concern over resources allocated to other people. Moreover, such preferences may be affected by the environment in which such choices are made, especially social networks and social pressure. I performed a lab-in-the-field experiment in rural China, where I recruited 162 Chinese farmers to vote in 7 variants of allocation games in randomly assigned groups and with real-world social contacts, with and without pressure. I find that social network and social pressure combined have significant yet heterogeneous effects on social preferences. The source of heterogeneity includes the assignment with in-group or out-group members, membership in dominant lineages, individual characteristics as defined by age and gender, and the degree of kinship between individuals within a social group. My study not only provides empirical evidence for the social preference theories but also urges policy makers to be careful in choosing an appropriate voting method. In addition, constraining the power of dominant lineage and having better educated villagers more involved in village affairs could be welfare improving.

BIOGRAPHICAL SKETCH

Xin Gao is a PhD student in Applied Economics and Management at Cornell University. Her primary fields are Development Economics, Labor Economics, and Behavioral/ Experimental Economics. The majority of her research employs the experimental approach (Lab/Lab-in-the-Field Experiments and Choice Experiments) to examine challenges faced by developing countries in Asia, especially rural public goods provision, labor market inequality, and human capital investment.

Prior to her doctoral studies, she worked as an analyst with NERA Economic Consulting in Tokyo, Japan. She obtained her Master's degree in Public Administration from Cornell University in 2011 and her two Bachelor's degrees from Peking University (Bachelor of Laws) in China and Waseda University (Bachelor of Arts) in Japan in 2009. She worked as a summer intern with the Asian Development Bank (ADB) in 2018.

For Mama.

ACKNOWLEDGEMENTS

I would like to thank the Chair of my dissertation committee, Professor Calum Turvey, for his guidance throughout my graduate years. I was lucky to have your input during our many meetings and to have your encouragement through the difficult times.

I would also like to express my deepest gratitude to the Co-Chair of my committee, Professor Nancy Chau. You are a great mentor who inspires my intellectual pursuit, a close friend whom I can always confide in, and most importantly a role model – the intelligent, elegant, passionate, and genuine woman that I always aspire to become one day. You gave me hope in a world like this.

My two minor members, Professor William Schulze and Professor Ricardo Daziano, provided me with the most precious guidance on the experimental design and estimation techniques of my work. This dissertation would not have been possible without their patience and expertise.

My sincere thanks are due to my two co-authors: Professor Hong Fu from the Shandong University of Finance and Economics, who managed and implemented the field experiments with me in China; and Aiko Kikkawa Takenaka from the Asian Development Bank (ADB), who served as my mentor during my internship in 2018. I have thoroughly enjoyed working with you.

I was blessed to be surrounded by intelligent and dedicated classmates, colleagues, and professors in the Dyson School as well as in the Department of Economics. Thank you for your suggestions and feedback on my work both in my seminars and in private.

Lastly, and most importantly, I would like to express my love and gratitude to my teachers and friends at the Ithaca Ballet. You are my home and my family. Thank you for making my journey less lonely and extraordinarily beautiful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The heart of the standard or rational model of economics is the idea that consumers seek to maximize innate, stable preferences over the quantities and attributes of the commodities they consume [1]. Along with the progress in choice theories, empirical techniques for the estimation of preference parameters have evolved from cross-sectional regressions using market-level or national-accounts-level data in the 1960s, to hedonic models using individual-level data, to the development of random-utility discrete-choice framework (McFadden 1968, 1976, 1978; Train 2003).

While choice analysis and preference estimation saw rapid growth in modeling consumer decisions and marketing strategies, its application in Development Economics, Labor Economics, and Public Economics has been scarce. However, it must be acknowledged that individuals, households, and communities in fact do have preferences over attributes of consumer goods as well as other goods, for example hedonic aspects of jobs, the value of education, and the fairness of resource allocation within a society. Just as consumer preference lies at the core of product development and market growth, preferences and choices over education and social goods are at the core of social development, since they can be both the causes and the consequences of inequality.

This dissertation therefore aims to examine how choices at the individual, household and community levels are shaped by preferences as well as the social context in which the decisions are made, and how they contribute to a variety of issues faced by developing countries in Asia, for example labor market inequality,

[1] McFadden, D. (2001). Economic choices. American economic review, 91(3), 351-378.

under-investment in human capital, and under-provision of rural public goods.

The first chapter, "*Labor Market Outcomes with Heterogeneous Preferences and Search Frictions: The Case of Chinese Migrant and Urban Workers*", examines decision-making at the individual level. The study originates from my observation from the Chinese labor market where white-collar jobs tend to be dominated by urban workers whereas blue-collar jobs with undesirable working conditions are usually undertaken by migrant workers. Yet migrant workers are paid less than urban workers, which is contradictory to Rosens classical compensating differentials model. In the theoretical part of my paper, I developed a search model with market frictions to show that wages need not be compensating when preferences are heterogeneous, and that the group more averse to undesirable working conditions need not earn less when reservation utilities differ and/or when employers practice taste-based discrimination. In the empirical part, I substantiated the assumption that urban workers are more averse to undesirable working conditions using a Discrete Choice Experiment, where I recruited 225 workers in China to make hypothetical choices between jobs characterized by different wage levels and working conditions. I backed out preference parameters and willingness-to-accept (WTA) measures for job attributes using multinomial logit and mixed logit regressions, both in the preference space and in the willingness-to-pay space. I find that consistent with my assumption, urban workers need to be compensated more to accept outdoor jobs and jobs in second-line cities. I also find that migrant workers have more dispersed preferences that vary with personal characteristics such as gender and education. The policy implication is that the implementation of labor contract law and workplace safety regulations may have unintended effects, in that it may induce urban workers into the market and crowd out migrant workers.

The second chapter focuses on households as decision makers. "*Where Did the Money and Time go? De-mystifying the Negative impact of Remittances on Human Capital Investment in the Kyrgyz Republic*", was developed during my internship with the Asian Development Bank (ADB) in the summer of 2018. We examine how households allocation of financial and labor resources affects human capital formation in the Kyrgyz Republic. It is generally believed that international migration and remittances from overseas encourage human capital investment and improve educational outcomes in developing countries. Empirical studies, however, have shown mixed evidence some positive, some negative, some have shown zero effects. In our study, we focus on the case of Kyrgyz Republic, one of the largest remittances-receiving countries in Asia. We used a 5-year panel dataset that tracks the same 3,000 households and 8,000 individuals in the country to examine the impact of remittances on both household educational expenditure and attendance rate of school age children. We used instrumental variables and fixed-effects regressions to correct for potential selection bias. We find that remittances have a negative impact on human capital formation namely both educational expenditure and school attendance rate are lower for households that receive a higher amount of remittances. To explore the possible channels of the negative effects, we further regressed itemized household expenditures and the time use pattern of school-age children on remittances. We find that the negative effects can at least be partly attributed to increased expenditure on durable goods and increased hours of child labor on farm work as a compensation for adult labor insufficiency induced by out-migration. Our finding calls for the need for monitoring farm labor hours of school-age children. Implementation and scaling up of financial literacy programs that that help parents balance short-term expenditures (durable goods) and long-term investments (education, health) can be beneficial.

The final chapter goes beyond individuals and households to examine communities and how social preferences and social network can affect collective decision-making."*Kinship, Social Preferences and Voting in Rural China: A Lab-in-the-Field Experiment*", is my job market paper. It is motivated by the policy change in rural China where public goods provision is no longer determined by the government but has to be voted by villagers. The implementation, however, has encountered many difficulties due to the unique social network structure where villages are shaped by lineages with conflicting interests and heterogeneous power. To understand how network structure and peer pressure affect villagers choices and behaviors in public goods voting, I performed a lab-in-the-field experiment in rural China, where I recruited 162 Chinese farmers along with their social contacts to vote in 7 variants of allocation games. The payoff structure reflects trade-offs among self-interest, efficiency, and the equality of allocation. The experiment follows a 2-by-2 design. On one dimension, I introduced "social network treatment by making the control group vote in randomly assigned 3-person groups while the treatment group votes in 3-person groups consisting one pair of social contacts plus one stranger. On the "social pressure dimension, the control group is assured that their votes will not be revealed at any point whereas the treatment group is informed beforehand that their votes will be made public to their group members at the end of the experiment. I used a multinomial logit model to back out preference parameters. I have three findings. First of all, I find that social network and social pressure combined have significant effects on social preferences. In particular, it makes the pairs more concerned with in-group gains at the cost of the out-group member. Secondly, social preference patterns vary with personal characteristics. For example, female subjects are more averse to inequality whereas more educated subjects care more about efficiency. Thirdly, the heterogeneity in preferences can also be explained by network characteristics. Namely, pairs with closer social relationships tend to focus more on within-group gains under pressure. My study not only rectifies the drawbacks of lab experiments in validating social preference theories but also has policy implications. In the setting of public goods referenda, policy makers need to choose appropriate voting methods, for example secret ballot instead of show of hands. Moreover, having female and more educated villagers more involved in village affairs may help improve public goods voting outcomes in rural China.

Although the entire dissertation falls into the broad area of Development Economics, the three chapters focus on Labor Economics, Labor/Education Economics, and Social/Public Economics respectively. In terms of methodology, the first chapter employs a combination of theoretical modeling and discrete choice experiment (DCE), the second chapter uses a more traditional fixed-effects panel regression with instrumental variables approach, whereas the third chapter is an application of lab-in-the-field experiment. In terms of geographical coverage, the first and third chapters focus on rural China while the second chapter examines the Kyrgyz Republic. The broad coverage of fields and techniques serves as an additional evidence that choices and preferences matter in multiple aspects of human existence and can be approached from multiple perspectives. The impetus for all the above endeavors, however, is my desire to understand why people make the choices they make, and how such choices reflect the human conditions they are faced with, and how such choices in turn contribute to the evolution of the human society.

CHAPTER 2

LABOR MARKET OUTCOMES WITH HETEROGENEOUS PREFERENCES AND SEARCH FRICTIONS: THE CASE OF CHINESE MIGRANT AND URBAN WORKERS

Abstract

Neither Rosens classical compensating differential model nor newly developed search models could explain the particular pattern of wage and job characteristics distributions in China, where migrant and urban workers coexist and dominate different sectors. In the theoretical part of this paper, I expand the model by Lang and Majumdar (2003) to show that wages need not be compensating when preferences are heterogeneous, and that the group more averse to undesirable working conditions need not earn less when reservation utilities differ and/or when employers practice taste-based discrimination. In the empirical part, I substantiate the assumption in my theoretical modeling that urban workers are more averse to undesirable working conditions using a discrete choice experiment, where 225 workers in China made hypothetical choices between jobs characterized by different wage levels and working conditions. I backed out preference parameters and willingness to accept measures for job attributes. I find that consistent with my assumption, urban workers need to be compensated more to accept outdoor jobs and jobs in second line cities. I also find that migrant workers have more dispersed preferences that vary with personal characteristics such as gender and education.

2.1 Introduction

The relationship between wage and non-pecuniary job characteristics such as working hours, working condition, contract and insurance offers has been explored theoretically and tested empirically for decades. The baseline theoretical model can be dated back to 1974, where Rosens equalizing differential, or compensating differential model, predicts that jobs offering unfavorable working conditions must pay premiums to attract workers. This prediction, however, is contradicted by numerous empirical tests[1]. Lang and Majumdar (2003) showed that the contradiction is likely to be due to the competitive market assumption. They incorporated search frictions into the model to show that in the presence of search frictions, the equilibrium wages need not be compensating. They then extend the model to allow for heterogeneous preferences and showed that when firms can make type-contingent offers, there will be wage discrimination against the group that is more averse to unattractive job characteristics. This prediction, however, is not consistent with observations in labor markets where multiple types of workers coexist. The most eminent case is China, where rural migrants often take jobs with undesirable characteristics that urban workers are unwilling to take (Zhao 2000; West and Zhao 2000; Meng 2000; Meng and Manning 2010). For example, according to a national representative survey of over 5000 urban and migrant workers in 2009 (RUMiC), 12.8% of urban workers and 77.7% of migrant workers are employed in blue collar positions, yet migrant workers not only earn less in terms of average monthly wage

[1]Lucas (1977) use a cross-sectional data source to estimate how wages varied with job characteristics. Although some of the coefficients for other job characteristics are of the expected sign, many others take exactly different signs. For example, jobs requiring physical strength are generally associated with lower wages than sedentary ones. Browns (1980) use a panel data analysis to exclude individual level unobservables, but still obtained many wrong-signed estimates relative to what is predicted by the theory. He concludes that the hypothesis that wrong signs in previous studies was due to the omission of important dimensions of worker quality was not supported by the data.

but also have lower access to insurance and fringe benefits compared to urban workers (Figure 2.1 and Figure 2.2). Our more up-to-date, albeit less geographically representative survey in 2017 in Shandong Province, reveals similar patterns (Figure 2.3). The average monthly wage for migrant workers in our sample is 6330.86 RMB, which is 10,000 RMB lower than their urban counterparts. Meanwhile, migrant workers face significantly worse working conditions (longer working hours and tougher working environment) and have lower rate of contract and insurance compared to urban workers.

A by sector comparison from our recent survey (Figure 2.4) offers more insights. In both white-collar and blue-collar jobs, migrant workers face significantly worse working conditions – longer hours, more dirty environment, less rate of contract and insurance coverage. Wage-wise, however, migrant workers on average earn significantly less in white-collar jobs and significantly more in blue-collar jobs.

Figure 2.1: Wage and Working Conditions : Urban Workers

VARIABLES	Mean	Standard Deviation	Min	Max
Sector Choice				
Blue collar	12.8%			
White collar	55.0%			
Unemployed	32.2%			
Labor Supply				
Hours supplied per year	1,381	1,020	0	4,320
hours supplied per day	5.753	4.252	0	18
Job Characteristics				
Monthly wage	1,977	3,046	0	80,000
Hourly wage	19.63	31.60	0	627.5
Have unemployment insurance	0.275	0.335	0	1
Have pension	0.306	0.327	0	1
Have injury insurance	0.283	0.385	0	1
Have housing fund	0.232	0.298	0	1
Personal Characteristics				
Age	47.61	9.903	20.95	64.99
Number of Kids	1.102	0.509	0	6
Year of Education	11.40	3.391	0	35
City Hukou	97.2%			
Non-Agriculture Hukou	97.5%			
Non-labor income	74,623	127,148	0	2.550e+06

Figure 2.2: Wage and Working Conditions : Migrant Workers

VARIABLES	Mean	Standard Deviation	Min	Max
Sector Choice				
Blue collar	77.7%			
White collar	21.5%			
Unemployed	0.7%			
Labor Supply				
Months in city per year	11.11	2.215	1	12
Hours supplied per year	2,521	826.2	0	5,376
Hours supplied per day	9.187	1.886	0	16
Job Characteristics				
Monthly wage	1,591	730.6	0	10,000
Hourly wage	7.433	4.336	0	75
Have unemployment insurance	0.0990	0.261	0	1
Have pension	0.145	0.296	0	1
Have injury insurance	0.146	0.313	0	1
Have housing fund	0.0539	0.192	0	1
Personal Characteristics				
Age	29.30	10.39	16	72
Number of Kids	0.646	0.919	0	9
Year of Education	9.526	2.843	0	17
City Hukou	16.3%			
Non-Agriculture Hukou	0.7%			
Non-labor income	1,050	6,355	0	132,000

Figure 2.3: Migrant-Urban Comparison: Wage and Working Conditions

	Urban	**Migrant**	**Difference**
Wage	16822.29	6330.86	-1.0e+04***
Hours	8.8	9.4	0.639***
Day	5.2	5.8	0.538***
Labourer	25.3%	26.8%	0.0150
Outdoor	19.3%	31.0%	0.117***
Danger	6.0%	9.9%	0.038***
Health Threat	8.4%	11.3%	0.028**
Heat	8.4%	14.8%	0.064***
Night Shift	12.0%	14.8%	0.027**
Standing	15.7%	26.1%	0.104***
Dirt	2.4%	11.3%	0.089***
Unemployment Insurance	47.0%	29.2%	-0.178***
Pension	48.8%	30.6%	-0.182***
Injury Insurance	52.4%	29.6%	-0.225***
Housing Fund	39.8%	15.5%	-0.243***
Contract	80.7%	56.0%	-0.247***

Figure 2.4: By Sector Comparison: Wage and Working Conditions

	White			Blue		
	Urban	*Migrant*	*Difference*	*Urban*	*Migrant*	*Difference*
Wage	25460.00	7653.00	-1.8e+04***	4157.41	5008.71	851.307***
Hours	8.01	8.73	0.721***	10.40	10.03	-0.367*
Day	5.27	5.45	0.175***	5.02	6.14	0.970***
Labourer	6.0%	12.5%	0.065***	66.7%	41.4%	-0.252***
Outdoor	6.0%	16.7%	0.107***	48.1%	45.7%	-0.0240
Danger	2.0%	1.4%	-0.00600	14.8%	18.6%	0.0380
Health Threat	4.0%	8.3%	0.043***	18.5%	14.3%	-0.042*
Heat	2.0%	5.6%	0.036***	22.2%	24.3%	0.0210
Night Shift	8.0%	5.6%	-0.024*	22.2%	24.3%	0.0210
Standing	22.0%	20.8%	-0.0120	7.4%	31.4%	0.240***
Dirt	2.0%	4.2%	0.022**	3.7%	18.6%	0.149***
Unemployment Insurance	45.0%	40.3%	-0.047**	57.4%	17.9%	-0.396***
Pension	50.0%	42.4%	-0.076***	53.7%	18.6%	-0.351***
Injury Insurance	54.0%	39.4%	-0.144***	57.4%	19.6%	-0.374***
Housing Fund	39.0%	20.1%	-0.189***	48.1%	10.7%	-0.374***
Contract	80.0%	59.2%	-0.208***	92.6%	52.9%	-0.397***

Another interesting observation in China is that the unemployment rate among urban workers is substantially higher. In fact, in recent years urban youths non-participation has become such a social problem that an abbreviation is created –"NEET" (Not currently engaged in Employment, Education or Training). The unemployment rate for migrants in the RUMiC 2009 sample was 1%, while for their urban counterparts was 37%. Our recent survey, although less representative, offers a similar pattern. All the migrants we interviewed are currently employed, while 7% of the urban workers are unemployed.

Figure 2.5: By Sector Comparison: Employment Rate

	Urban	**Migrant**
Percent in White	60.24%	50.70%
Percent in Blue	32.53%	49.30%
Percent Unemployed	7.23%	0.00%

Therefore, the above evidence from China contradicts both Rosen and Lang & Majumdar's predictions. In the theoretical part of this paper, we employ and extend Lang and Majumdars model to explain the unemployment rate gap between urban and migrant workers in China. We also extend their model to show that when both reservation utility and aversion to undesirable job characteristics differ, the group that is less averse to undesirable working conditions could be paid less. Last but not least, we explore the impact of taste-based discrimination – either when both sectors practice taste-based discrimination or only the white-collar sector discriminates against migrant workers – on labor market outcomes of both populations. In the empirical part, we conduct a discrete choice experiment (DCE) with migrant workers and urban workers in China to back out preference parameters for each population. This allows us to robustly test preference heterogeneity and calculate willingness to accept (WTA) measures of each group. This helps us substantiate the assumption in the theoretical model that urban workers are more averse to undesirable working conditions than migrant workers.

This study contributes to the existing literature by modifying Lang and Majumdars model to explain the empirical reality in the labor markets of many developing countries. Our discrete choice experiment contributes to the empirical labor economics literature by isolating the effect of preferences from market constraints, which allows us to robustly estimate preference parameters and willingness to accept measures. The rest of this paper is organized as follows. Section 2.2 reviews existing literature both in theoretical modeling and empirical estimation. In Section 2.3 we extend Lang and Majumdars model to explain the wage and unemployment gaps and incorporate taste-based discrimination into the model. In Section 2.4 we present the design and results of our discrete choice experiment (DCE) in China. Section 2.5 concludes.

2.2 Literature Review

2.2.1 Theoretical Modeling of Compensating Differentials

Labor market participation, or working, can no longer be viewed as merely a means to earning money and feeding oneself. Jobs must be viewed multidimensionally, where non-pecuniary characteristics such as working hours, working conditions, locations, contract and insurance offers, are as important as monetary payoffs. Workers choose jobs based on not only wage offers but also the consumption values of the jobs, thereby generating an implicit market for job characteristics. Firms must balance the prospect of hiring and the costs to improve attractiveness of job offers. The market equilibrium and the wage and welfare distributions resulting from the above process are of great interest to labor economists.

Rosens (1974, 1986) equalizing differential, or compensating differential model, offers a baseline analysis. When labor market is competitive, jobs that offer favorable working conditions attract labor at lower than average wages, whereas jobs offering unfavorable working conditions must pay premiums to attract workers. If two jobs, one clean and one dirty, is offered, then the market must only observe two wages, with the difference W1-W0 just enough to compensate the marginal worker to make her indifferent between choosing the two jobs.

The compensating differential, however, is hardly supported by empirical works. Lucas (1977) use a cross-sectional data source to estimate how wages varied with job characteristics. Although some of the coefficients for other job characteristics are of the expected sign, many others take exactly different signs. For example, jobs requiring physical strength are generally associated with lower wages than sedentary ones. Browns (1980) use a panel data analysis to exclude individual level unobservables, but still obtained many wrong-signed estimates relative to what is predicted by the theory. He concludes that the hypothesis that wrong signs in previous studies was due to the omission of important dimensions of worker quality was not supported by the data.

The contradiction between theory and data is very likely to be because the assumption of competitive market in Rosens model is not realistic. Lang and Majumdar (2003) incorporated search frictions into the model to show that in the presence of frictions in the labor market, the equilibrium job distribution need not show evidence of compensating wage differentials. They then extend the model to allow for two types of workers with different preferences for non-pecuniary job characteristics and conclude that when firms can make type-contingent offers, there will be wage discrimination against the group that is more averse to unattractive job characteristics. Namely, if urban workers are more averse to dirty work than migrant workers, then urban workers will on average earn less than migrant workers in both clean and dirty jobs.

2.2.2 Wage Distributions and Occupational Segregation in China

The urban labor market in China in the past decade has witnessed a sharp increase in the inflow of migrant workers from rural area. From the late 1990s, the number of rural migrants increased by more than 100 million and another 300 million inflow is expected for the next few decades. An interesting phenomenon is that China's rural migrants often take jobs which urban workers are unwilling to take (Zhao

2000; West and Zhao 2000; Meng 2000; Meng and Manning 2010). It is estimated that in 2011, over 89 percent of migrant workers are employed as unskilled workers in blue-collar jobs (construction, manufacturing, etc.), while only 40 percent of urban workers take these jobs (Meng 2012). Blue-collar jobs in urban China are characterized by low wage, long working hours, undesirable working conditions, and inaccessibility to contract and social insurance. For example, the average monthly wage of migrant workers in 2009 is 1,591 RMB, 400 RMB lower than their urban counterparts, although they on average work 1,100 hours per year more than urban workers . It is also reported that as of 2010, the proportion of migrant workers with access to unemployment insurance is 13.5%, while the proportion for workers with urban hukou is 66%. The proportion of migrants with access to urban health insurance is 20% in 2010, while for urban hukou workers it is 87% (Frijters, Gregory, and Meng 2015).

One can conjecture that such occupational exclusion is due to the difference in skill distribution, namely urban residents are almost exclusively skilled labors and are fully absorbed by white-collar jobs. A contradiction, however, is that the unemployment rate among urban hukou workers is substantially high. The unemployment rate for migrants in 2009 was 6%, while for their urban hukou counterparts in the same cities it was 37% (Meng 2012). This implies that 37% of urban residents would rather stay at home and live off their unemployment benefits than take on blue collar jobs. In fact, urban youths non-participation has become so prominent a phenomenon that an abbreviation is created – "NEET" (Not currently engaged in Employment, Education or Training). It is therefore doubtful that the occupational exclusion between the migrants and urban workers can be fully explained by skill distribution and discrimination differences in preferences must have been playing an important part in sorting and selection.

On the policy side, the Chinese government has implemented multiple policies to improve the welfare of migrant workers. In 2008, China introduced a new Labor Contract Law (LCL) to protect workers rights. The majority of scholars that examined the impact of the law (Cheng et al., 2015; Cui et al., 2013; Gallagher et al., 2014; Li & Freeman, 2015, Meng, 2017) found positive effects on labor market outcomes such as wages, working hours, and other work related benefits. Others find that it reduced labor market flexibility and pushed up wages and labor costs due to the macro-economic downturn in recent years (Luo, 2015). While the LCL requires employers to specify working conditions in the written contract, it does not provide penalty for employers who do not pay the insurance premium. In recent years, the government has been contemplating to establish a nationwide unified social security account system that makes workers social insurance accounts transferable from one locality to another (Li, 2008). Predicting the potential impact of this new policy on the welfare of migrants and urban workers will therefore offer valuable policy insights.

Such policy evaluations and simulations can hardly be achieved using reduced form approaches. One widely used approach in the Chinese labor literature is hedonic regression using cross-sectional and longitudinal data (Kong and Wang 2012, 2014), building on the theoretical framework of equalizing differences in Rosen (1974) and Rosen (1986). However, estimates from this approach are unstable to adding person or workplace controls and are often wrong-signed due to the fact that individuals self-select into jobs that have certain attributes correlated with wage levels, as well as measurement error and the presence of search frictions in the labor market (Hwang et al., 1998; Lang and Majumdar, 2004; Bonhomme and Jolivet, 2009). Additionally, reduced-form parameters are only combinations of primitive parameters and are of little use in performing counterfactual analysis and policy simulations. Our study therefore contributes to this literature by providing a more accurate and unbiased estimation of migrants and urban workers valuations for written contracts, insurance, and workplace safety. Since the new policies, once implemented, will be available to both migrants and urban workers, we use their valuations to simulate how the two populations would respond to the new policies and predict how the employment share of each group in each occupation would change.

2.2.3 Discrete Choice Experiments and Preference Parameters

Isolating the effects of preferential choices and constraints on occupational exclusion has always been a challenging task in labor economics (Altonji and Blank, 1999). Labor economists who studied occupational segregation by gender, for example, disagree on whether differences in job characteristics between the jobs held by men and women should be counted as constraints that women face in the labor market or as an indication of differential tastes by women for the jobs that they want to hold (flexibility for example). It is also argued that group differences in pre-labor market human capital investment and in non-labor market activities may lead to differences in comparative advantage across occupations. Meng (2012) uses a linear probability regression to predict whether an individual has a white-collar or blue-collar job and finds that after controlling for all observable individual and market characteristics, migrants are still around 16 to 24 percent less likely to have a white-collar job. This indicates that over and above their attribute differences there is still a great portion of unexplained factors that contributes to the difference in occupational choices. In addition, the reduced form approach that Meng employs does not take self-selection into participation into account and is likely to yield biased estimates. Unfortunately, although structural estimation (Adda et al 2017) and experiments (Benjamin et al 2010, Flory et al. 2014) have been widely applied to the study preferential differences by gender and race, efforts to comparing preferential choices by migrants and urban workers in China have been scarce, if not non-existing.

Eliciting stated preference (SP) with choice experiments has been widely used in the analyses of consumer choices (cars, transportation modes, health care etc.). It is not until recent years that labor economists start to employ this approach. Eriksson and Kristensen (2014), for example, use a vignette method to elicit WTP for various job amenities and fringe benefits in an internet sample of Dutch respondents. Wiswall and Zafar (2016) use a stated preference approach to study how undergraduate students value job characteristics such as availability of part-time work and potential for promotion in hypothetical future jobs. Mas and Pallais (2017) use a discrete choice experiment in the employment process for a national call center to estimate the willingness to pay distribution for alternative work arrangements such as flexible working hours and working from home. The advantage of these approaches is that they robustly identify preferences for various job attributes, free from omitted variable bias and free from considering the equilibrium matching of workers to jobs (Wiswall and Zafar, 2016). The disadvantage to the approach is that it is unclear to what extent responses to hypothetical questions are accurate and approximate behavior in a market setting. This concern has led to a large literature probing hypothetical bias in the context of contingent valuation surveys (see e.g., List and Shogren, 1998).

The current study is to our knowledge the first attempt to perform choice experiment in the Chinese labor market. We conduct a discrete choice experiment (DCE) with migrant workers and urban hukou workers in China to robustly estimate and directly compare parameters that characterize preferences among these two populations.

2.3 Theoretical Framework

We start our analysis with a theoretical model, where we show that preference parameters play a key role in determining wage distribution and unemployment incidences. We omit the simplified homogeneous workers part in Lang and Majumdar's (2003) and start with two types of workers (Migrant and Urban) with heterogeneous preferences both searching for jobs on a market with frictions.

2.3.1 Firm and Profit Function

Each job offer has two dimensions: Salary S and working condition h, with higher h indicating worse working condition. Final product $Q(h)$ is increasing in h with price normalized to 1. Firms incur a cost $C(h)$ to improve working condition and $C'(h) < 0$. Firm's profit function is therefore

$$\pi = Q(h) - S - C(h)$$

2.3.2 Worker and Utility Function

There are x types of workers. For now we assume $x = 2$: urban (U) and rural (R). Workers differ only in Hukou status and are equally productive. They both like higher wage ($\partial U/\partial S > 0$) and dislike bad working conditions ($\partial U/\partial h < 0$), but to different extents. For now we assume urban workers are more averse to bad working conditions compared to rural migrants – an assumption that will be tested with a choice experiment. Utility function is assumed to be additively separable in wage and working condition, :

$$U_x(S,h) = \Phi(S) - \beta_x \cdot h \quad j = U, R$$

$$\beta_U > \beta_R > 0$$

where $\Phi'(S) > 0$ and $\Phi''(S) < 0$. We also assume that compared to rural migrants, urban workers have higher reservation utility. That is

$$U_U^r > U_R^r > 0$$

2.3.3 Market Frictions Equilibrium without Taste-based Discrimination

We start out with a competitive labor market with search frictions where employers do not practice taste-based discrimination. The number of firms with job vacancies for type x workers is M_x. The number of type x workers looking for jobs is N_x. The vacancy-worker ratio $\lambda_x = \frac{M_x}{N_x}$ measures market tightness and follows a Poisson distribution. In this section we assume that N_x is large. Each firm decides independently whether or not to make an offer, and how best to offer it. We assume that firms are aware of the Hukou status of the workers and can make type-contingent offers. Making an offer will incur a fixed cost $K > 0$. Each worker chooses among all offers he receives the one that gives him the highest utility (which has to be no lower than his reservation utility). Since offer arrival is random, some workers receive zero offers and will be unemployed. Similarly, some firms' offers will be turned down and their vacancies will remain unfilled.

Firms face the trade-off between making a better offer to increase the probability of its offer being accepted, and the fact that better offers lower profit. Here a firm's mixed strategy is a probability distributions over all offer bundles (S, h) and a probability of making no offer. The firm's expected profit is

$$E_x(\pi) = P_x(U_x(S, h))(Q(h) - S_x - C(h)) - K$$

where $P_x(\cdot)$ is the probability of the offer being accepted by a type $x = U, R$ worker. Since the worker chooses the offer that gives him the highest utility, only utility ranking matters. The following two claims are quoted from Lang and Majumdar (2003) without proving.

Claim 1: In any equilibrium, the offer distribution must be continuous, have support $[U_x^r, U_x^{max}]$, and have no mass points.

Claim 2: Define $\tilde{\pi}_x(a) = max_{\{S,h:U_x(S,h)=a\}}Q(h) - S - C(h)$, and the utility level U_x^{max} by $\tilde{\pi}_x(U_x^{max}) = K$, then if $e^{-\lambda_x}\tilde{\pi}_x(U_x^r) < K$, all offers in the support of any equilibrium distribution must make zero expected profit. All offers that are outside the support must make non-positive expected profit.

Intuitively, Claim 1 establishes a continuous distribution bounded from below by the reservation utilities of each group (since any offer lower than reservation utility will not be accepted) and from above by the maximum utility that each group can achieve (since firms will not make an offer that generates negative profit). $\tilde{\pi}_x(a)$ in Claim 2 establishes the most profitable way to offer a bundle of salary S and working condition h that ensures $U_x(S, h) = a$. Under mild concavity conditions [2], the bundle of (S, h) offer that ensures $U_x(\cdot) = a$ is unique. Similarly, U_x^{max} defines the maximum achievable utility for type x worker when firms make zero expected profit.

Employment Incidences

For simplicity, instead of taking h as a continuum of working conditions, we assume for now that firms can only offer one of the two types of working conditions: Blue-collar ($h = 1$) or White-collar ($h = 0$). Claim 3 is cited from Lang and Majumdar (2003) without proof.

Claim 3: If dirty jobs are more profitable when the lowest possible wages are paid:

$$Q(1) - C(1) - \Phi^{-1}(U_x^r + \beta_x) > Q(0) - C(0) - \Phi^{-1}(U_x^r)$$

then the lowest utility offer for group x specifies a dirty job; If workers prefer clean jobs when the highest possible wages are paid:

$$\Phi(Q(0) - C(0) - K) > \Phi(Q(1) - C(1) - K) - \beta_x$$

then the highest utility offer for group x specifies a clean job; If both conditions are met, then all offers above some cut-off utility level U_x^* specify that $h = 0$ and all offers below U_x^* specify that $h = 1$. U_x^* is given by the equation:

$$Q(1) - C(1) - \Phi^{-1}(U_x^* + \beta_x) = Q(0) - C(0) - \Phi^{-1}(U_x^*)$$

U_x^* is non-decreasing with respect to $[Q(1) - C(1)] - [Q(0) - C(0)]$, and $U_R^* > U_U^*$.

The two conditions ensure that both blue-collar and white-collar jobs exist in equilibrium.

[2] $U_s Q''(h) + U_{hh} + 2U_{Sh}Q'(h) + U_{SS}(Q'(h))^2 < 0$, which is satisfied, for example, when both the utility funchion $U_x(\cdot)$ and the profit function $Q(h) - C(h) - S$ are concave

Proposition 1: Unemployment rate for urban workers is higher.

Proof: According to Claims 1& 2, the equilibrium with market frictions must involve mixed strategies and firms must make zero expected profit. Namely,

$$\begin{aligned} P_x(U_x(S,h)) &= \frac{K}{Q(h) - C(h) - S_x} \\ &= \frac{K}{Q(h) - C(h) - \Phi^{-1}(U_x + \beta_x \cdot h)} \end{aligned}$$

$P_x(U_x^r)$, namely the cumulative density of utility offer measured at reservation utility, gives the unemployment rate of each group. According to Claim 3, the lowest utility offers come from blue-collar jobs ($h = 1$), we only need to look at the lower end of the blue-collar offers to compare unemployment rate:

$$\begin{aligned} P_R(U_R^r(S,1)) &\equiv \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R)} \\ &< \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)} \equiv P_U(U_U^r(S,1)) \end{aligned}$$

Namely, the unemployment rate for urban workers is always higher since they have higher reservation utility and stronger distaste for bad working conditions, namely $U_U^r + \beta_U > U_R^r + \beta_R$.

Proposition 2: In equilibrium, the proportion of migrant workers employed in blue-collar jobs is higher than that of urban workers. The proportion of urban workers employed in white-collar jobs is higher than that of migrant workers.

Proof: Based on Claim 3, the proportion of each type of worker in blue-collar jobs is characterized by the cumulative density below cut-off utility and above reservation utility. Namely,

$$\begin{aligned}
P_U(h=1) &= P_U(U_U^*) - P_U(U_U^r + \beta_U) \\
&= \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^* + \beta_U)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)} \\
&= \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_U^*)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)}
\end{aligned}$$

$$\begin{aligned}
P_R(h=1) &= P_R(U_R^*) - P_R(U_R^r + \beta_R) \\
&= \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^* + \beta_R)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R)} \\
&= \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_R^*)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R)}
\end{aligned}$$

Since $U_U^* < U_R^*$ according to Claim 3, $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)} > \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}$. Also since $U_U^r > U_R^r$ and $\beta_U > \beta_R$, we have $\frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)} < \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}$. Therefore $P_U(h=1) < P_R(h=1)$, which means at equilibrium a higher proportion of migrant workers will be hired in blue-collar jobs compared to urban workers.

Similarly, the proportion of each type of worker in white-collar jobs is therefore characterized by the cumulative density above the cut-off utility. Namely,

$$\begin{aligned}
P_U(h=0) &= 1 - \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_U^*)} \\
P_R(h=0) &= 1 - \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_R^*)}
\end{aligned}$$

Since $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)} > \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}$, we have $P_U(h=0) > P_R(h=0)$, which indicates at equilibrium, a higher proportion of urban workers will be hired as white-collars than migrants.

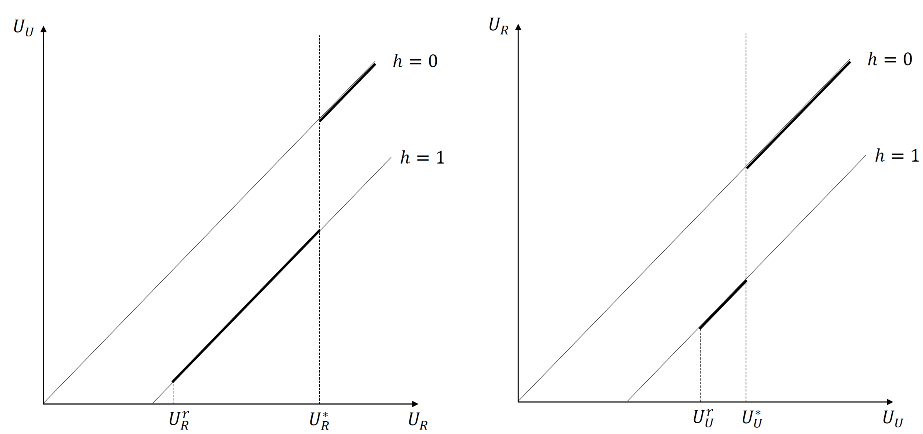
Wage Incidences

In this sub-section we look at how wage distributions would be affected by market frictions and preference heterogeneity.

Proposition 3: In a frictional competitive labor market where employers do not practice taste-based discrimination, the average wage of migrant workers is higher in white-collar jobs. The average wage of migrant workers may be higher or lower in blue-collar jobs.

Proof: We first characterize 4 key points: (1) the highest wage offers achievable in white-collar jobs for both types; (2) the lowest wage levels offered in white-collar jobs to both types; (3) the highest wage offers achievable in blue-collar jobs for both types; (4) the lowest wage levels offered in blue-collar jobs to both types. These points combined with the share of each population employed in each sector will portray an intuitive explanation for average wage levels, as illustrated in Figure 2.6. We then more rigorously characterize wage distributions in terms of stochastic dominance.

Figure 2.6: Offer Distributions for Migrant and Urban Workers (Without Discrimination)



(a) Migrant Workers (R) (b) Urban Workers (U)

According to Claim 2, the maximum level of utility that both types of workers can achieve is when $\tilde{\pi}(U_x^{max}) = K$, namely when firms make zero profit. It is intuitive that

$$U_U^{max} = U_R^{max} \equiv \tilde{\pi}^{-1}(K)$$

In white-collar jobs we have $h = 0$ so neither type of workers needs to be compensated for disutilities from undesirable working conditionss. Therefore the highest wage levels that both types of workers can expect are equal. Namely,

$$\bar{S}_{R,h=0} \equiv \Phi^{-1}(U_R^{max}) = Q(0) - C(0) - K = \Phi^{-1}(U_U^{max}) \equiv \bar{S}_{U,h=0}$$

The lowest wage offered in white-collar jobs are characterized by wages at the cut-off utility levels. According to Claim 3, $U_R^* > U_U^*$, therefore we have:

$$\underline{S}_{R,h=0} \equiv \Phi^{-1}(U_R^*) > \Phi^{-1}(U_U^*) \equiv \underline{S}_{U,h=0}$$

It is therefore clear that although the proportion of migrant workers employed in white-collar jobs is smaller compared to urban workers, their average wage is higher in white-collar jobs. This can be mathematically shown by comparing the distributions of salaries S conditional on the worker being employed in the white sector ($h = 0$) for the two types of workers. For white collar jobs, at any given utility level u,

$$\begin{aligned}
&P_U(S \leq \Phi^{-1}(u)|h = 0) - P_R(S \leq \Phi^{-1}(u)|h = 0) \\
&= \frac{\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}}{1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}} - \frac{\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)}}{1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)}} \\
&= \frac{(1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(u)})(\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)})}{(1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)})(1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)})} > 0
\end{aligned}$$

Namely, $P_U(S \leq \Phi^{-1}(u)|h = 0) > P_R(S \leq \Phi^{-1}(u)|h = 0)$ which indicates that in white collar jobs, the salary distribution for migrant workers first-order stochastically dominates that for urban workers. Therefore, the average wage for migrant workers is higher in white collar jobs.

For blue-collar jobs, highest utility offers would be at the compensated cut-off utilities for the two types of workers, namely $\bar{S}_{U,h=1} \equiv \Phi^{-1}(U_U^* + \beta_U)$ and $\bar{S}_{R,h=1} \equiv$

$\Phi^{-1}(U_R^* + \beta_R)$. By the definition of U_x^* in Claim 3, $Q(1) - C(1) - \Phi^{-1}(U_x^* + \beta_x) = Q(0) - C(0) - \Phi^{-1}(U_x^*)$. Therefore,

$$\bar{S}_{U,h=1} \equiv \Phi^{-1}(U_U^* + \beta_U) = [Q(1) - C(1)] - [Q(0) - C(0)] + \Phi^{-1}(U_U^*)$$

$$\bar{S}_{R,h=1} \equiv \Phi^{-1}(U_R^* + \beta_R) = [Q(1) - C(1)] - [Q(0) - C(0)] + \Phi^{-1}(U_R^*)$$

Since $U_R^* > U_U^*$, we have

$$\bar{S}_{R,h=1} - \bar{S}_{U,h=1} = \Phi^{-1}(U_R^*) - \Phi^{-1}(U_U^*) > 0$$

Which means the highest wage offer achievable is higher for migrant workers than for urban workers in the blue-collar sector.

The lowest wage offers to both groups in the blue-collar sector are characterized by wage offer at reservation utilities. By assumption, $U_R^r < U_U^r$ and $\beta_R < \beta_U$, we have

$$\underline{S}_{R,h=1} \equiv \Phi^{-1}(U_R^r + \beta_R) < \Phi^{-1}(U_U^r + \beta_U) \equiv \underline{S}_{U,h=1}$$

This indicates that the lowest wage offer to migrant workers is lower compared to urban workers in blue-collar jobs, which is opposite to the upper ends of the wage offers across the two groups.

Therefore, intuitively, the average wage of migrant workers in blue collar jobs may or may not be higher than that of urban workers, depending on the relative size of the differences at both ends and the proportion of workers employed in each sector. We can also use similar method to compare the salary distribution for both types of workers conditional on the worker being employed as a blue-collar worker.

$$P_U(S \leq \Phi^{-1}(u + \beta_U)|h = 1) = \frac{\frac{K}{Q(0)-C(0)-S} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}}{\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}}$$

$$P_R(S \leq \Phi^{-1}(u + \beta_R)|h = 1) = \frac{\frac{K}{Q(0)-C(0)-S} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)}}{\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)}}$$

Since both the numerator and denominator are larger for migrant workers ($P_R(S \leq \Phi^{-1}(u + \beta_R)|h = 1)$), we cannot ascertain the relative size of the two. Therefore migrant workers may earn higher or lower wage on average in blue-collar jobs.

Our result is different from Lang and Majumdar (2003), where they claimed that the type of worker more averse to undesirable working conditions will have lower average wage in both sectors. This is because they assumed that both types of workers have the same reservation wage – which in our case would indicate that urban workers' reservation utility would be lower than that of migrant workers. By assuming that urban workers have higher reservation utility, our model adds a lower tail to the wage distribution of urban workers, thereby bringing down the average wage of urban workers.

2.3.4 Market Equilibrium When Both Sectors Practice Taste-based Discrimination

The employment incidences in the previous section coincide with the reality in China, in that urban residents have higher unemployment rate and are more likely to be employed in white-collar jobs. The wage incidences, however, are contrary to what we observe in our most recent survey data. In this section, we allow employers to perform taste-based discrimination against migrant workers. In particular, firms generate negative utility from hiring a migrant worker which translates into a monetary cost, $d > 0$. This can be thought of as a reduction in productivity of other workers since they have to work with someone that they do not want to work with. The profit function when firms in both sectors practice taste-based discrimination can be written as

$$E_x(\pi) = P_x(U_x(\cdot))(Q(h) - S - C(h) - d \cdot \mathbb{1}_{x=R}) - K$$

where $\mathbb{1}_{x=R}$ is an indicator function that takes value 1 if the worker is migrant. Workers' utility function remain the same as in the previous section.

Employment Incidences

We first explore the employment incidences when firms of both sectors practice taste-based discrimination.

Proposition 4: Unemployment rate of migrant workers will increase when both sectors practice taste-based discrimination. It may be higher or lower than that of urban workers, depending on the size of discrimination d. If the size of discrimination is greater than the difference in compensating differentials between the two groups, that is $d > \Phi^{-1}(U_R^r + \beta_U) - \Phi^{-1}(U_R^r + \beta_R)$, then the unemployment rate of migrant workers will exceed that of the urban workers.

Proof: Since Claims 1 & 2 still hold in this scenario, firms must make zero expected profit at equilibrium. Namely,

$$\begin{aligned} P_x(U_x(S,h)) &= \frac{K}{Q(h) - C(h) - S_x - d \cdot \mathbb{1}_{x=R}} \\ &= \frac{K}{Q(h) - C(h) - \Phi^{-1}(U_x + \beta_x \cdot h) - d \cdot \mathbb{1}_{x=R}} \end{aligned}$$

Similar to the previous section, unemployment rate is determined by the cumulative density at reservation utility of each group, that is, $P_x(U_x^r)$.

$$P_R(U_R^r(S,1)) \equiv \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R) - d}$$

$$P_U(U_U^r(S,1)) \equiv \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)}$$

The inclusion of d in the denominator for the calculation of migrant unemployment increases the size of $P_R(U_R^r(S,1))$, indicating that when migrant workers are discriminated against, their unemployment rate would rise. However, since $\Phi^{-1}(U_R^r + \beta_R) < \Phi^{-1}(U_U^r + \beta_U)$, we cannot compare the relative size of $\frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)-d}$ and $\frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}$, hence the relative size of $P_R(U_R^r(S,1))$ and $P_U(U_U^r(S,1))$ cannot be ascertained. The key takeaway here, however, is that when both sectors practice taste-based discrimination, migrant workers may face higher unemployment rate than urban workers despite their lower aversion to bad working conditions and lower reservation utility.

We then derive and compare the proportions of the two groups employed in blue-collar and white-collar jobs.

Proposition 5: When employers in both sectors perform taste-based discrimination against migrant workers, the proportion of migrant workers employed in white-collar jobs will be smaller than that of urban workers. Furthermore, it decreases from the level where firms do not discrimination. The proportion of migrant workers hired in blue-collar sector may be higher or lower than that of urban workers, depending on the size of discrimination d. If the size of discrimination is greater than the difference in compensating differentials between the two groups, that is $d > \Phi^{-1}(U_R^r + \beta_U) - \Phi^{-1}(U_R^r + \beta_R)$, then the proportion of migrant workers employed in the blue-collar sector will be lower than that of the urban workers.

Proof: Since the inclusion of d in firms' profit functions does not change the concavity conditions, Claim 3 can be carried over here. The cut-off utilities U_x^* where firms are indifferent between offering type x worker a blue-collar or white-collar takes the same form as in Claim 3, except that the profit functions when migrant workers are hired involve an additional term d. The cut-off points for R

and U are now defined as:

$$Q(1) - C(1) - \Phi^{-1}(U_R^* + \beta_R) - d = Q(0) - C(0) - \Phi^{-1}(U_R^*) - d$$

$$Q(1) - C(1) - \Phi^{-1}(U_U^* + \beta_U) = Q(0) - C(0) - \Phi^{-1}(U_U^*)$$

Since d is subtracted from both side of the first equation, the cut-off utilities stay exactly the same as when firms do not practice taste-based discrimination. The cut-off utilities for urban workers also stayed the same as the second equation is unchanged. As a result, the conclusion in Claim 3 remains valid in this case. Namely,U_x^* is non-decreasing with respect to $[Q(1) - C(1)] - [Q(0) - C(0)]$ and $U_R^* > U_U^*$.

When taste-based discrimination is present, the proportions of migrant and rural workers employed in white-collar jobs are defined by the line segments above the cuf-off utilities, which can be written as:

$$P_U(h = 0) = 1 - \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_U^*)}$$

$$P_R(h = 0) = 1 - \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_R^*) - d}$$

since $U_R^* > U_U^*$ and $d > 0$, we have $\Phi^{-1}(U_U^*) < \Phi^{-1}(U_R^*) + d$ and hence $P_U(h = 0) > P_R(h = 0)$. Namely, the proportion of urban workers employed in white-collar jobs is higher than that of migrant workers. Further more, the subtraction of d from the denominator decreases the size of $P_R(h = 0)$ from the level where taste-based discrimination did not exist.

Similarly, the proportions of migrant and rural workers employed in blue-collar jobs are defined by the line segments above the reservation utilities and below the cut-off utilities, which can be written as:

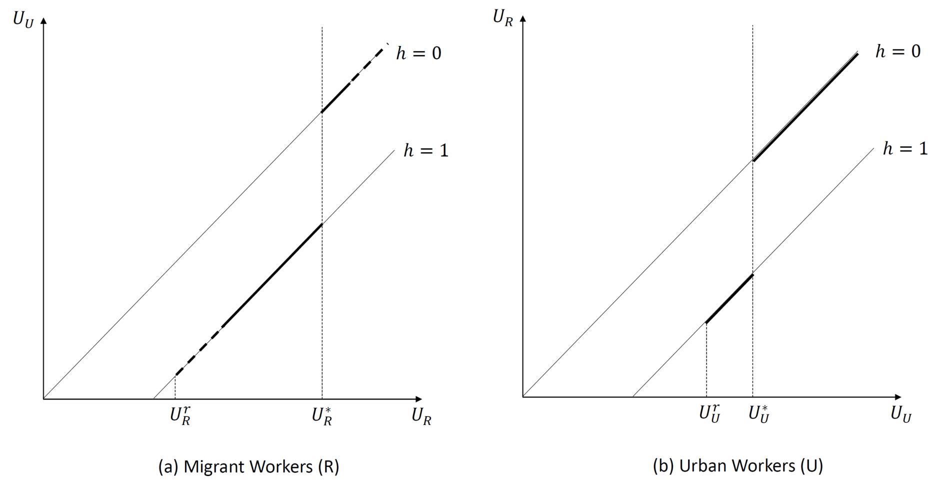
$$\begin{aligned}
P_U(h=1) &= P_U(U_U^* + \beta_U) - P_U(U_U^r + \beta_U) \\
&= \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^* + \beta_U)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)} \\
&= \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_U^*)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)}
\end{aligned}$$

$$\begin{aligned}
P_R(h=1) &= P_R(U_R^* + \beta_R) - P_R(U_R^r + \beta_R) \\
&= \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^* + \beta_R) - d} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R) - d} \\
&= \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_R^*) - d} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R) - d}
\end{aligned}$$

We have shown in the white-collar case that $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)-d} > \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}$. The relative size of $\frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}$ and $\frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)-d}$, however, cannot be ascertained because it depends on whether $\Phi^{-1}(U_R^r + \beta_R) + d$ is greater than or less than or equal to $\Phi^{-1}(U_U^r + \beta_U)$. Therefore, the proportion of migrant workers employed in blue-collar jobs may or may not be higher than that of urban workers.

The offer distributions for migrant and urban workers when both sectors discriminate are illustrated in Figure 2.7. The cut-off utilities (U_x^*) remain unchanged. However, since both sectors generate a disutility from hiring a migrant worker, the upper tail of the offer distribution in white-collar sector will render a negative profit for the firms. Similarly, offering the reservation utility in blue-collar jobs will end up having a negative profit. Hence both line segments – above as well as below the cut-off utility – grew shorter in this scenario.

Figure 2.7: Offer Distributions for Migrant and Urban Workers (Both Sectors Discriminate)



Wage Incidences

As employment incidences change when firms practice taste-based discrimination, wage distribution also change. While in the non-discrimination case, migrant workers are predicted to have higher average wage compared to urban workers in white-collar jobs, we show in this section that this is not necessarily the case.

Proposition 6: The average wage of migrant workers may or may not be higher than that of urban workers, regardless of sectors.

Proof: We first look at the highest achievable wages for both populations in white-collar jobs. Similar to the proofs before, firms make zero profit when the highest utility levels are offered. That is, $\tilde{\pi}_x(U_x^{max}) = K$. However, different from the previous sections, $\tilde{\pi}_x(\cdot)$ takes different form for migrant workers and rural workers now, due to the presence of discrimination. For migrant workers, maximum achievable utility solves

$$Q(0) - \Phi^{-1}(U_R^{max}) - C(0) - d = K$$

while for urban workers, it solves

$$Q(0) - \Phi^{-1}(U_U^{max}) - C(0) = K$$

Consequently, wage offer at the top point of the distribution is lower for migrant workers than for urban workers. That is

$$\Phi^{-1}(U_R^{max}) \equiv Q(0) - C(0) - d - K < Q(0) - C(0) - K \equiv \Phi^{-1}(U_U^{max})$$

Namely, $\bar{S}_{R,h=0} < \bar{S}_{U,h=0}$. Intuitively, since firms perceive the presence of migrant workers as a negative impact on overall productivity, they would not offer as good a package to migrant workers as in the non-discrimination scenario.

The lowest achievable wage in the white-collar sector is, similar to the non-discrimination scenario, determined by the cut-off utilities. We discussed in the employment coincidence section that the cut-off points remain the same as in the non-discrimination scenario when both sectors practice discrimination against the migrant population. By Claim 3, the cut-off utility for migrant workers is higher than for urban workers, the lowest wage offered to migrants is higher than that to urban workers. Namely:

$$\underline{S}_{R,h=0} \equiv \Phi^{-1}(U_R^*) > \Phi^{-1}(U_U^*) \equiv \underline{S}_{U,h=0}$$

Since we have $\bar{S}_{R,h=0} < \bar{S}_{U,h=0}$ and $\underline{S}_{R,h=0} > \underline{S}_{U,h=0}$, it is impossible to ascertain if the average wage of migrant workers is higher or lower than that of urban workers. The relationship depends on the relative size of discrimination d and the difference between the two populations' degrees of aversion to undesirable job characteristics. Using distributional expressions, we can show that the probability that a worker receives a wage offer lower than S conditional on the offer being a white-collar job, for each type of worker respectively, is:

$$P_U(S \leq \Phi^{-1}(u)|h=0) = \frac{\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}}{1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}}$$

$$P_R(S \leq \Phi^{-1}(u)|h=0) = \frac{\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)-d} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)-d}}{1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)-d}}$$

Although P_U has a larger denominator than P_R, we cannot compare the relative size of the numerators since $\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)} < \frac{K}{Q(0)-C(0)-\Phi^{-1}(u)-d}$ and $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)} < \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)-d}$. One thing we can ascertain from this practice though is that when discrimination is present, migrant workers will not necessarily earn more on average in white-collar jobs than urban workers do.

The comparison between average wages in blue-collar sector follows the same logic. Since the cut-off utilities did not change, the highest achievable wage is higher for migrants than for urban workers in blue-collar sector, that is

$$\begin{aligned}\bar{S}_{R,h=1} &\equiv \Phi^{-1}(U_R^*) + [Q(1) - C(1)] - [Q(0) - C(0)] \\ &> \Phi^{-1}(U_U^*) + [Q(1) - C(1)] - [Q(0) - C(0)] \equiv \bar{S}_{U,h=1}\end{aligned}$$

The lowest wage level offered should be lowered when discrimination is present in blue-collar jobs. However, as no packages than reservation utility will be accepted, the lowest wage offer is still at reservation level. Since $U_R^r < U_U^r$ and $\beta_R < \beta_U$, we have

$$\underline{S}_{R,h=0} \equiv \Phi^{-1}(U_R^r + \beta_R) < \Phi^{-1}(U_U^r + \beta_U) \equiv \underline{S}_{U,h=1}$$

Since$\bar{S}_{R,h=1} > \bar{S}_{U,h=1}$ and $\underline{S}_{R,h=1} < \underline{S}_{U,h=1}$, again we cannot determine if migrant workers earn more or less than urban workers on average in blue-collar jobs. This is the same as the case of non-discrimination. Alternatively, the probability that a worker receives a wage offer lower than S conditional on the offer being a blue-collar job, for each type of worker respectively, is:

$$P_U(S \le \Phi^{-1}(u+\beta_U)|h=1) = \frac{\frac{K}{Q(0)-C(0)-S} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}}{\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}}$$

$$P_R(S \le \Phi^{-1}(u+\beta_R)|h=1) = \frac{\frac{K}{Q(0)-C(0)-S-d} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)-d}}{\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)-d} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)-d}}$$

Again it is impossible to determine the relative size of P_U and P_R, because P_R has both larger numerators and denominators than P_U. Therefore, when both sectors practice taste-based discrimination against migrant workers, wage incidences in both sectors become unclear. Migrant workers may or may not earn higher wage in either sector.

2.3.5 Market Equilibrium When Only White Sector Discriminates

In this section, we explore the distributional consequences of employment and wage when only white-collar employers perform taste-based discrimination. Namely, only firms that offer white-collar jobs experience a reduction in profit $d > 0$ when a migrant worker is hired. This is a more realistic description of the current urban job market in China, since working environments such as construction sites and laundry shops are less likely to generate disutilities from hiring migrant workers than office jobs. In the case of white-collar job discrimination alone, the profit function can be modified to incorporate another indicator function $\mathbb{1}_{h=0}$ that takes value 1 if the job is a white-collar one ($h = 0$)

$$E_x(\pi) = P_x(U_x(\cdot))(Q(h) - S - C(h) - d \cdot \mathbb{1}_{x=R} \cdot \mathbb{1}_{h=0}) - K$$

Same as the previous case, $\mathbb{1}_{x=R}$ is an indicator function that takes value 1 if the worker is migrant. Workers' utility function remain the same as in the previous section.

Employment Incidences

We first explore the employment incidences.

Proposition 7: The unemployment incidence returns to the non-discrimination level, with the unemployment rate of urban workers higher than that of migrant workers.

Proof: The unemployment rates for both sections, as shown in the proof for Proposition 1, is still measured by $P_x(U_x^r)$, namely cumulative distribution of utility offer measured at reservation utility. Since blue sector does not discriminate against migrant workers, the unemployment rate returns to the non-discrimination level, namely:

$$
\begin{aligned}
P_R(U_R^r(S,1)) &\equiv \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R)} \\
&< \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)} \equiv P_U(U_U^r(S,1))
\end{aligned}
$$

The unemployment rate for urban workers in blue-collar jobs is always higher since they have higher reservation utility and stronger distaste for bad working conditions, namely $U_U^r + \beta_U > U_R^r + \beta_R$.

Proposition 8: The proportion of migrant workers employed in white-collar jobs is smaller than that of urban workers. Moreover, this proportion is not only smaller than the one in non-discrimination scenario but also smaller than the on in the case where both sectors discriminate against migrant workers.

Proof: We first show that when only white sector practices taste-based discrimination, the cut-off utility for migrant workers (denoted as $U_R^{*'}$ here) no longer remains at the cut-off level of non-discrimination and the case where both sectors discriminate (U_R^*). This is because as per Claim 3, cut-off utility level is where employer is indifferent between offering a blue-collar job and a white-collar job:

$$Q(1) - C(1) - \Phi^{-1}(U_R^{*'} + \beta_R) = Q(0) - C(0) - \Phi^{-1}(U_R^{*'}) - d$$

Since $U_R^{*'}$ is non-decreasing in $[Q(1) - C(1)] - [Q(0) - C(0) - d]$, we have $U_R^{*'} > U_R^*$. Intuitively, when white-collar jobs suffer from a disutility of hiring a migrant worker, employers find it more efficient to shift earlier to offering a blue-collar job. The cut-off utility for urban workers remains at the same level ($U_U^{*'} = U_U^*$), since the profit functions remain the same in both blue and white sectors.

The proportion of each type of workers hired in white-collar jobs is therefore characterized by the cumulative density above the cut-off utility. Namely,

$$P_U(h = 0) = 1 - \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_U^{*'})}$$
$$P_R(h = 0) = 1 - \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_R^{*'}) - d}$$

Since $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^*)-d} > \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}$, we have $P_U(h = 0) > P_R(h = 0)$. That is, a higher proportion of urban workers will be hired as white-collars than migrants.

The proportion of migrants hired as white-collar workers when only white section practices taste-based discrimination is not only lower than the non-discrimination case but also lower than the case where both sectors discriminate against migrants. The combination of two forces led to this result. First of all, the presence of discrimination moved the upper end of the offer distribution since the original highest utility offer U_R^{max} will yield negative profit

$Q(0) - C(0) - \Phi_{-1}(u^{max}) - d < K$ so it will no longer be offered. Secondly, the fact that only white-collar jobs see disutility from employing a migrant worker makes employers want to shift to offering a blue-collar job earlier, rendering the lower end of the white-collar offer distribution cut short as well.

Proposition 9: When only white-collar employers discriminate against migrant workers, the proportion of migrants employed in blue-collar jobs is higher than that of urban workers. Furthermore, it is higher than the case where both sectors practice taste-based discrimination.

Proof: Similar to previous sections, the proportion of each type of worker in blue-collar jobs is characterized by the cumulative density below cut-off utility and above reservation utility. Namely,

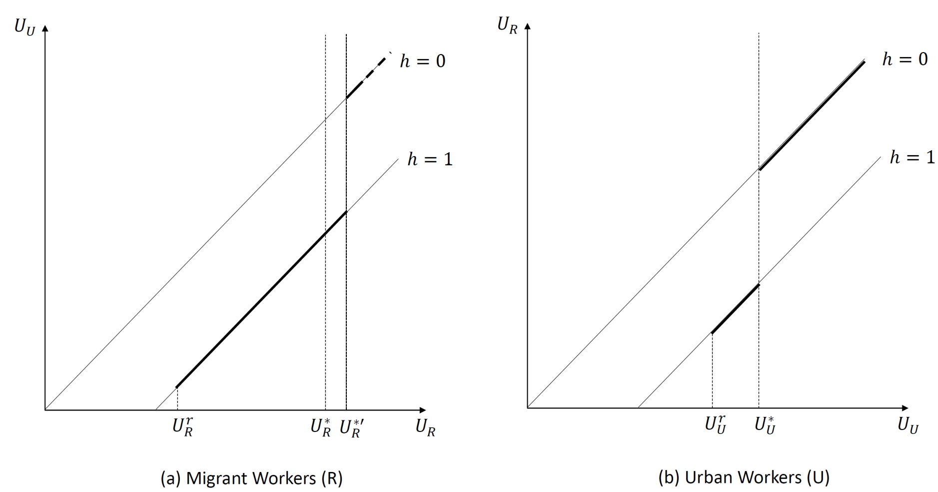
$$\begin{aligned}
P_U(h=1) &= P_U(U_U^*) - P_U(U_U^r + \beta_U) \\
&= \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^* + \beta_U)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)} \\
&= \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_U^*)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_U^r + \beta_U)}
\end{aligned}$$

$$\begin{aligned}
P_R(h=1) &= P_R(U_R^{*'}) - P_R(U_R^r + \beta_R) \\
&= \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^{*'} + \beta_R)} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R)} \\
&= \frac{K}{Q(0) - C(0) - \Phi^{-1}(U_R^{*'}) - d} - \frac{K}{Q(1) - C(1) - \Phi^{-1}(U_R^r + \beta_R)}
\end{aligned}$$

Since $U_U^* < U_R^* < U_R^{*'}$, $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^{*'})} > \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}$. Also since $U_U^r > U_R^r$ and $\beta_U > \beta_R$, we have $\frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)} < \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}$. Therefore $P_U(h=1) < P_R(h=1)$, which means at equilibrium a higher proportion of migrant workers will be hired in blue-collar jobs compared to urban workers. In addition, the proportion of migrants hired in blue collar in this case also exceeds the non-discrimination case and the case where both sectors practice discrimination since $U_R^{*\prime} > U_R^*$. In summary, when only whit-collar employers discriminate migrant workers, the migrants' unemployment rate would not increase as in the case where both sectors discriminate. However, even less of them will be hired in the white sector and more would end up in blue sector.

The offer distributions for migrant and urban workers in this scenario are illustrated in Figure 2.8. Cut-off utility for migrant workers is now higher than the previous two scenarios, since firms find it more profitable to shift to a blue-collar job earlier. The line segment above U_R^*, therefore, is not only shorter than the baseline scenario but also shorter than the scenario where both sectors discriminate. Since firms do not generate a negative utility from offering a blue-collar job to migrant workers, the lower end of the offer distribution in blue-collar sector returns to the baseline scenario level.

Figure 2.8: Offer Distributions for Migrant and Urban Workers (One Sector Discriminates)



Wage Incidences

Proposition 10: The average wage of migrant workers may or may not be higher than that of urban workers, regardless of sectors. However, migrant workers are better-off in terms of wage distribution compared to the case where both sectors discriminate.

Proof: The highest achievable wages for both populations in white-collar jobs stay at the same level as the case where both sectors discriminate against migrant workers. Namely, for migrant workers, maximum achievable utility solves

$$Q(0) - \Phi^{-1}(U_R^{max}) - C(0) - d = K$$

while for urban workers, it solves

$$Q(0) - \Phi^{-1}(U_U^{max}) - C(0) = K$$

and

$$\Phi^{-1}(U_R^{max}) \equiv Q(0) - C(0) - d - K < Q(0) - C(0) - K \equiv \Phi^{-1}(U_U^{max})$$

Namely, $\bar{S}_{R,h=0} < \bar{S}_{U,h=0}$. When we look at offers in the white-collar sector, the top ends stay the same as in the case where both sectors discriminate because employers are following the same practice.

The lowest achievable wage in the white-collar sector is determined by the cut-off utilities. As shown in the proof of Proposition 8, $U_R^{*'} > U_R^* > U_U^*$, we have

$$\underline{S'}_{R,h=0} \equiv \Phi^{-1}(U_R^{*'}) > \Phi^{-1}(U_R^*) > \Phi^{-1}(U_U^*) \equiv \underline{S}_{U,h=0}$$

That is, the new lowest achievable wage of migrant workers is not only higher than that of urban workers but also higher than the case where both sectors discriminate.

Similar to the previous case, since we have $\bar{S}_{R,h=0} < \bar{S}_{U,h=0}$ and $\underline{S'}_{R,h=0} > \underline{S}_{U,h=0}$, it is impossible to ascertain if the average wage of migrant workers is higher or lower than that of urban workers. The relationship again depends on the relative size of discrimination d and the relative degree of aversion. It is interesting to notice, however, that although a smaller portion of migrant workers are employed in the white sector, their average wage does improve compared to the case where both sectors discriminate since the lower end of their wage distribution is removed.

This can also be more rigorously shown by calculating the probability that a worker receives a wage offer lower than S conditional on the offer being a white-collar job when only the white-collar sector practices taste based discrimination:

$$P_U(S \leq \Phi^{-1}(u)|h=0) = \frac{\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}}{1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)}}$$

$$P_R(S \leq \Phi^{-1}(u)|h=0) = \frac{\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)-d} - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^{*'})-d}}{1 - \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^{*'})-d}}$$

Very similar to the case where both sectors discriminate, although P_U has a larger denominator than P'_R, we cannot compare the relative size of the numerators since $\frac{K}{Q(0)-C(0)-\Phi^{-1}(u)} < \frac{K}{Q(0)-C(0)-\Phi^{-1}(u)-d}$ and $\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)} < \frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^{*'})-d}$. Therefore, we cannot tell if migrant workers on average are offered higher or lower wage than urban workers. However, since $U_R^{*'} > U_R^*$, we can show that

$$P_R(S \leq \Phi^{-1}(u)|h=0) - P'_R(S \leq \Phi^{-1}(u)|h=0) > 0$$

That is, in this case, migrant workers do see an improvement in their average wage offer compared to the case where both sectors discriminate.

Wage distributions in blue-collar sector also changed due to the shift of cut-off utility points. We can see that now the highest achievable wage of migrant workers is not only higher than that of urban workers, but also higher than the migrant workers' highest achievable wage in the case where both sectors discriminate:

$$\begin{aligned}\bar{S}'_{R,h=1} &\equiv \Phi^{-1}(U_R^{*\prime}) + [Q(1) - C(1)] - [Q(0) - C(0)] + d \\ &> \Phi^{-1}(U_U^{*}) + [Q(1) - C(1)] - [Q(0) - C(0)] \equiv \bar{S}_{U,h=1}\end{aligned}$$

The lowest wage level offered, similar to the previous two cases, will be determined by reservation utilities. Since $U_R^r < U_U^r$ and $\beta_R < \beta_U$, we have

$$\underline{S}_{R,h=0} \equiv \Phi^{-1}(U_R^r + \beta_R) < \Phi^{-1}(U_U^r + \beta_U) \equiv \underline{S}_{U,h=1}$$

Since$\bar{S}'_{R,h=1} > \bar{S}_{U,h=1}$ and $\underline{S}_{R,h=1} < \underline{S}_{U,h=1}$, again we cannot determine if migrant workers earn more or less than urban workers on average in blue-collar jobs. However, since the wage distribution of blue-collar workers now contains an upper tail that was not present when both sectors discriminate, migrant workers are on average offered a higher wage in this case. This can also be more rigorously shown using the probability distribution of wages conditional on the offer being a blue-collar job:

$$P_U(S \leq \Phi^{-1}(u + \beta_U)|h = 1) = \frac{\frac{K}{Q(0)-C(0)-S} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}}{\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_U^*)} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_U^r+\beta_U)}}$$

$$P'_R(S \leq \Phi^{-1}(u + \beta_R)|h = 1) = \frac{\frac{K}{Q(0)-C(0)-S-d} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)}}{\frac{K}{Q(0)-C(0)-\Phi^{-1}(U_R^{*\prime})-d} - \frac{K}{Q(1)-C(1)-\Phi^{-1}(U_R^r+\beta_R)}}$$

Since P'_R has both larger numerators and denominators than P_U, it is impossible to determine the relative size of the two. Therefore, as in the case where both sectors practice taste-based discrimination migrant workers may or may not earn higher wage in either sector when only the white sector discriminates. However, it is evident that migrant workers are better-off in terms of average wage offer when only one sector discriminates, since $U_R^{*\prime} > U_R^*$,

$$P_R(S \leq \Phi^{-1}(u)|h = 1) - P'_R(S \leq \Phi^{-1}(u)|h = 1) > 0$$

That is, migrant workers' average wage offer in the blue-collar sector improves compared to the case where both sectors discriminate.

To summarize, in the theoretical part of this paper, we derived that in the presence of market search frictions and heterogeneous preferences, compensating differentials predicted by Rosen's classical model may not be present. Our model also explained why the unemployment rate of urban workers is higher than that of migrant workers. Moreover, contradicting the prediction of Lang and Majumdar (2003), when the group more averse to undesirable working conditions (i.e. urban workers) has higher reservation utilities, they do not necessarily earn less in both sectors. We also show that when employers practice taste-based discrimination, the group less averse to undesirable working conditions (i.e. migrant workers) may lose their wage advantage in both sectors. Moreover, if both sectors discriminate against migrant workers, the unemployment rate of migrant workers may in fact exceed that of urban workers.

2.4 Empirical Validation: The Discrete Choice Experiment

To substantiate our assumption that urban workers are more averse to undesirable working conditions, namely $\beta_U > \beta_R$, we implemented a Discrete Choice Experiment (DCE) where we randomly sampled 165 migrant workers and 93 urban workers in Shandong Province, China. The experiment allows us to obtain each group's stated preferences for wage and the non-monetary aspects of jobs, and calculate their willingness to accept (WTA) for undesirable job characteristics. The following section explains our experiment design, implementation, estimation strategy using the random utility framework, and results.

2.4.1 Experimental Design

The aim of the experiment is to back out each group of workers' preference parameters for wage and non-monetary aspects of jobs. We restrict the geographical coverage to the labor market of one location – Shandong Province of China. This is because the wage distribution differ greatly across provinces depending on the economic development level of the province. Workers at different locations, therefore, have very different expectations for wage and working conditions offered. If we were to cover a national representative sample, a more dispersed distribution of wage levels will need to be included in the choice experiment, which will increase the mental task of the respondents. It is therefore optimal to restrict the sampling to one local labor market. We chose Shandong because it is a middle-level province in terms of economic development. It is also a relatively large labor market that attracts a large number of migrant workers from all over the country.

The hypothetical jobs are constructed to best reflect the reality of contemporary labor market in Shandong. Each sampled worker responds to 4 hypothetical choice scenarios where in each scenario he has to consider the trade-off between the wage and attributes of two listed alternatives and choose his preferred one or opt out. This also allows us to estimate the willingness to accept (WTA) distribution of each group for undesirable job characteristics, namely, how much a worker needs to be compensated to accept a job with undesirable working conditions. Each respondent also fills out a questionnaire that collects basic demographic information after he completes the choice task. We combine information gathered from the questionnaire and the choice experiment to examine if within group heterogeneity in WTA can be explained by personal characteristics such as gender, age, education, and family income.

Constructing Attributes and Attribute Levels

The first step of designing this choice experiment is to decide what job attributes should be included in the portfolio and the range of variation of each attribute (the levels). Attributes included in the experiments should capture what choice makers deem relevant and important whereas the levels of attributes should reflect the reality and at the same time include a wide range of variation so that substitution patterns can be identified. To achieve these goals, we extract attributes and their distributions from both public data and a pilot survey we conducted in August 2016. The public data was extracted from the 2015 National Beural of Statistics Report[3]. We summarize job characteristics that are deemed important in the past literature by industry. We further implemented a pilot survey where we interviewed 30 migrant and urban workers in Shandong to collect information on their wage, working hours, and working conditions. Distributions of attributes are summarized in Figure 2.9. We observe that wage ranges roughly from 2000 RMB to 4500 RMB per month, and sectors with more disamenities (for example construction) offer relatively higher wages. Working hours range from 9 to 11 hours a day and 6-7 days a week. Although more than two-third of our surveyed migrants reported to have signed contract with their employers, less than half of them are enrolled in social insurance programs and more than half of them face danger.

[3]Source: Pilot survey in Shandong Province by authors, August 2016; Report on Chinese Migrant Workers 2015, National Bureau of Statistics of PRC

Figure 2.9: Attributes and Levels from Public Data and Survey

Industries	Percentage		Wage (RMB/Month)		Hours	Contract	Insurace	Danger
	Public	*Survey*	*Public*	*Survey*	*Survey*	*Survey*	*Survey*	*Survey*
Manufacture	31.1%	10.7%	2,970	3,333	9.0/ 6.3	66.7%	66.7%	66.7%
Construction	21.1%	71.4%	3,508	4,244	9.5/ 6.0	80.0%	35.0%	40.0%
Wholesale/ Retail	11.9%	-	2,716	-	-	-	-	-
Transportation/ Storage/Postal Services	6.4%	3.6%	3,553	1,800	10.0/ 7.0	100.0%	0.0%	10.0%
Accommodations/ Restaurants	5.8%	-	2,723	-	-	-	-	-
Residential Service/ Maintenance	10.6%	14.3%	2,686	4,325	11.0/6.3	75.0%	0.0%	100.0%

Note that the public data from National Bureau of Statistics of China uses a nation-wide sample, which may not accurately reflect wage levels in Shandong Province. We therefore conducted two focus group interviews one with migrant workers and the other with urban residents in Shandong – where we presented participants with the attribute levels and asked them if these are realistic and if they would like to add any other attributes that they deem important when they search for jobs. Almost all participants suggested that wage be adjusted upward. We therefore set 4 wage levels ranging from 3000 to 6000 RMB per month, which covers both the average wage range of migrants and urban workers. Some migrants suggested location whether the job is in large or small cities would matter. We therefore added location as an additional attribute and provided 3 levels – First-line cities, Second-line cities, Counties for migrants and First-line cities, Second-line cities, Third-line cities for urban workers (it is unrealistic for urban residents to go work in counties). The final construct of attributes and attributes levels are presented in Figure 2.10.

Figure 2.10: Attributes and Attribute Levels

Levels	Wage	Hours	Environment	Danger	Contract	Insurance	Location
1	3000	8hrs 5days	Office	Low	Yes	Yes	First Line
2	4000	9hrs 6days	Indoor but Not Office	Medium	No	No	Second Line
3	5000	10hrs 7days	Outdoor	High			Counties
4	6000	11hrs 7days					

Generating Choice Profiles

We use JMP to generate a Bayesian D-Optimal experimental design according to Sandor and Wedel (2001). This design criterion seeks to minimize the determinant of the variance-covariance matrix of the parameter estimators. Sandor and Wedel (2001) showed that the D-optimal designs generally outperform the linear design where prior parameter vector is set to be zero with zero prior variance (Huber and Zwerina 1996).

In our design, each respondent is presented with 4 choice scenarios where each choice scenario provides 2 alternative jobs that are presented as bundles of job attributes from Figure 2.10. Since it is widely acknowledged in choice modeling literature that "respondents often find it difficult to trade off prospective goods when every attribute of the offering changes in each comparison, especially in studies involving many attributes" (Kessels et al. 2011), we keep 3 attributes constant and vary 4 attributes at a time in each scenario. We have 16 choice scenarios in total that are divided into 4 versions of surveys. Each respondent is randomly assigned 1 version.

Each scenario begins with the following statement (in Chinese): "Imagine that you are currently not employed and you are actively looking for jobs. If these are the only two jobs you are offered and all other aspects that are not included in the table are identical across alternatives, which one will you choose?" Our prior mean parameters are set assuming that unattractive levels of an attribute can be compensated for by attractive levels of another attribute. To take into account the possibility that the above compensatory decision making assumption can be violated, we provide an opt-out option in each choice scenario where respondents can reject both alternatives and choose to stay unemployed for 6 months and keep looking for other jobs. Figure 2.11 shows a sample choice scenario that a respondent is presented with.

Figure 2.11: A Sample Choice Scenario

V1Q1

	Alternative 1	Alternative 2	Alternative 3 (Stay at home)
Wage	3000	5000	
Time	8 hours/day 5 days/week	8 hours/day 5 days/week	
Environment	Indoor, but not office	Outdoor	NA
Contract	Yes	No	
Insurance	Yes	Yes	
Danger	High	Low	
Location	Second-line	Second-line	Current City
Your Choice			

The choice experiment is followed by a questionnaire where we collect information on the respondents' demographic characteristics (gender, age, education, experience, etc.) as well as information on their current jobs (occupation, wage, hours, working conditions, access to contract and social insurance).

Experiment Implementation

Prior to the formal implementation of our experiment, we implemented two rounds of pilot experiment. The first round took place in December 2016 at Shandong University of Finance and Economics where 81 undergraduate and graduate students from 4-5 majors participated. The purpose was to test if the experiment instructions and the presentation is clear and easy to understand. The second round of piloting took place in February 2017, where we interviewed 50 migrant workers that returned to their hometown during Chinese New Year and 30 urban workers with low-medium level of education and non-elite jobs so that the migrant and urban samples are comparable. We also asked respondents for their feedback on the clarity and length of the survey and adjusted the format accordingly.

The formal experiment was implemented in July 2017 in Shandong Province in China in two stages. In Stage 1, we randomly sampled incoming passengers at the train station at Jinan City, a transportation hub where migrant workers come in for jobs and urban residents return home from either work or vacation. In Stage 2, we interviewed workers at construction sites, grocery stores and office buildings in Jinan City to obtain a balanced sample of blue-collar and white-collar workers. Since we do not know the *hukou* status (rural vs urban) a priori, the proportion of migrant and urban workers is not half-half – 142 migrant workers and 83 urban workers completed the choice experiment and the questionnaire. The demographic characteristics of the migrant sample and the urban sample are summarized in Table 2.1. The percentage of female, age, and number of kids are roughly balanced across samples. However, urban workers do have higher level of education and higher marriage rate compared to migrant workers.

Table 2.1: Summary Statistics for Migrant and Urban Samples

	Migrant	Urban	Difference
Female	50%	60.20%	-10.20%
Age	35.718	36.699	-0.98
Year of Education	11.93	14.41	-2.480***
Marriage	0.697	0.855	-0.158***
Number of Kids	0.923	0.964	-0.041
Number of Observations	142	83	

The experiment was conducted face-to-face on a one-on-one basis. The experimenter starts with introducing the purpose of the survey and asking the respondent if he or she would willingly participate. After the respondent agrees, the experimenter reads the experiment instructions on the questionnaire to the respondent (see Appendix). Each respondent is randomly assigned one version of the questionnaire that contains 4 choice scenarios. He or she is given 1 minute for each choice scenario to choose between the 2 alternative jobs or choose to opt out. In order to make sure the respondent understands the task and is paying attention, the experiment randomly pick one choice scenario and ask the respondent why he or she chose one alternative job over the other. After the choice experiment portion is finished, each respondent proceeds to fill out a short questionnaire on their demographics and current job situation. The whole process takes 8-10 minutes. Each respondent is thanked with a small gift equivalent to 8-10 RMB.

2.4.2 Estimation and Results

Our estimation is carried out in both the preference space and the willingness-to-pay (WTP) space. In the preference space estimation, we obtain both groups' preference parameters for wage as well as for each job characteristic in our experiment design (the *main effects*). We also test how such preferences vary with personal characteristics (the *interaction effects*). The drawback of the preference space estimation, despite its computational convenience, is that it often yields unreasonable willingness to pay estimates. To rectify this, we implement the WTP space estimation by directly specifying the distributions of WTP measures (although in our case it is the willingness to accept, or WTA).

Estimation in the Preference Space and Results

We start the preference space estimation with a standard Multinomial Logit (MNL) model, where respondents choices are analyzed in a random utility framework (McFadden 1974; Train 2003). Each respondent is faced with a set of alternatives and chooses the alternative that gives him or her the highest utility. Specifically, the utility of individual n choosing alternative i is:

$$U_{in} = ASC_i + q'_{in}(\beta + Z'_n\gamma) + \varepsilon_{in}$$

Where β and γ are main effect parameters and interaction effect parameters respectively. ASC_i is the alternative specific constant, q_{in} indicates a vector of job attributes ($K \times 1$, including wage) that individual n faces. Z is an $L \times 1$ vector of individual characteristics of person n, which is excluded from the baseline main effects estimations. ε_{in} is the unobserved part of the utility, which is assumed to follow an Extreme Value Type 1 distribution with mean 0 and variance 1. The probability that person n chooses alternative i is therefore:

$$P_{in} = Pr(U_{in} > U_{jn}) = \frac{e^{ASC_i + q'_{in}(\beta + Z'_n\gamma)}}{\sum_j e^{ASC_j + q'_{jn}(\beta + Z'_n\gamma)}}$$

Under the standard multinomial logit framework, parameters β and γ are estimated by maximizing the likelihood of sample joint choice probability:

$$l(\beta, \gamma; y|X) = \Pi_i^N \Pi_{j \in C_n} P_{in}^{y_{in}}$$

We divide the sample into two sub-samples – migrant workers and urban workers – to obtain preference parameters and interaction effects for each sub-sample respectively. The simple multinomial logit model, however, assumes that within each sub-sample everyone has the same taste and fails to take into account the possible heterogeneity within each group. To allow for taste heterogeneity, we estimate a

Mixed Logit model by assuming that the main effect parameters vector β follows a normal distribution parametrically characterized by θ (in this case, the mean and standard deviation of β_k). Namely

$$\beta_k \sim f(\beta_k|\theta_{\beta_k})$$

The reason why we assume the interaction effects parameters γ are non-random is to alleviate computational burden. The probability that individual n chooses alternative i in the mixed logit can be written as:

$$P_{in} = \int P_n(i|\beta_{k,n})f(\beta_k)d\beta_k = \int \frac{e^{ASC_i + q'_{in}(\beta + Z'_n\gamma)}}{\sum_j e^{ASC_j + q'_{jn}(\beta + Z'_n\gamma)}} f(\beta_k)d\beta_k$$

The distribution parameter vector θ is estimated by maximizing sample likelihood while accounting for the fact that each individual was faced with 4 choice situations ($T_n = 4$). This is to make sure that the heterogeneity estimated truly captures variations across individual, not within the same individual. The sample likelihood can therefore be written as:

$$L(\theta_{\beta_k}; y|X) = \Pi_{i=1}^{N} \int \Pi_{t=1}^{T_n} \Pi_{i \in C_n}(P_{int}|\varepsilon)f(\varepsilon)d\varepsilon$$

The Multinomial Logit estimation results for migrant workers, urban workers, as well as the difference between their preference parameters are presented in the table below.

Table 2.2: Main Effects by Sub-Group – Multinomial Logit

VARIABLES	(1) Migrant	(2) Urban	(3) Difference ((2)-(1))
Wage	0.000333***	0.000215***	-0.000118
	(5.69e-05)	(6.94e-05)	(8.97e-05)
Time	-0.0100***	-0.00535	0.00468
	(0.00382)	(0.00495)	(0.00625)
Outdoor	-0.422**	-1.170***	-0.747**
	(0.182)	(0.266)	(0.323)
Nonoffice	-0.316	-0.487	-0.170
	(0.221)	(0.315)	(0.385)
Contract	0.403***	0.276	-0.127
	(0.132)	(0.170)	(0.215)
Insurance	0.693***	0.690***	-0.00333
	(0.151)	(0.219)	(0.266)
MidDanger	-0.0151	-0.104	-0.0891
	(0.196)	(0.264)	(0.329)
HighDanger	-0.626***	-0.810***	-0.184
	(0.200)	(0.283)	(0.346)
FirstLine	-0.0548	0.0633	0.118
	(0.217)	(0.305)	(0.375)
SecondLine	0.382**	-0.142	-0.524*
	(0.169)	(0.239)	(0.293)

Standard errors in parentheses

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

As shown in the first two columns, both migrant workers and urban workers prefer higher wages and jobs that provide insurances, and dislike outdoor and highly dangerous jobs. The difference, as shown in column (3), is that compared to migrant workers, urban workers have a significantly stronger distaste for outdoor jobs and for jobs that are located in second-line cities. These results substantiate the assumption in our theoretical model that urban workers are more averse to undesirable working conditions compared to rural migrants, that is, $\beta_U > \beta_R > 0$.

Table 2.3: Main Effects by Sub-Group – Mixed Logit

	(1)	(2)	(3)	(4)
VARIABLES	Migrant (Mean)	Migrant (SD)	Urban (Mean)	Urban (SD)
Wage	0.000257*	-0.000547***	6.41e-05	0.000413**
	(0.000134)	(0.000132)	(0.000176)	(0.000184)
Time	-0.0363***	0.0453***	-0.0503***	0.0625***
	(0.0106)	(0.00979)	(0.0162)	(0.0148)
Outdoor	-0.755**	0.0652	-1.977***	-0.994
	(0.324)	(0.594)	(0.574)	(0.659)
Nonoffice	-0.848**	0.0333	-1.009	0.349
	(0.390)	(0.818)	(0.616)	(1.267)
Contract	0.420*	0.937***	-0.0216	-0.0386
	(0.235)	(0.316)	(0.278)	(0.673)
Insurance	0.864***	0.478	0.447	1.772**
	(0.270)	(0.710)	(0.429)	(0.811)
MidDanger	0.0747	1.064**	0.0456	-0.695
	(0.349)	(0.506)	(0.451)	(0.849)
HighDanger	-0.810**	-0.154	-0.728	-0.288
	(0.316)	(0.455)	(0.460)	(0.841)
FirstLine	-0.158	1.351**	0.00882	0.336
	(0.375)	(0.663)	(0.505)	(1.658)
SecondLine	0.491*	-0.712	-0.224	-0.0885
	(0.274)	(0.736)	(0.372)	(0.606)

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

Table 2.3 presents regression results from the Mixed Logit specification. Assuming that preference parameters follow a normal distribution within each subsample, we present the mean and standard deviation estimates separately. We observe that while the majority of standard deviation estimates from the migrant sample (column 2) are significant, those from the urban sample (column 4) are not. This indicates that migrant workers have more disperse preferences among themselves, while the preference pattern of the urban workers is more uniform.

Table 2.4 and Table 2.5 present the interaction effects between job characteristics and personal characteristics, namely, how preferences for job attributes vary with gender, age, education etc. A quick comparison between Table 2.4 (interaction effects for the migrant sample) and Table 2.5 (interaction effects for the urban sample) reveals that more parameters are statistically significant in Table 2.4, meaning that the variation in preferences are better explained by personal characteristics for the migrant sample than for the urban sample. This is consistent with our Mixed Logit estimation results, as the preference pattern of migrant workers are more dispersed.

Table 2.4: Interaction Effects (Migrant Workers)

VARIABLES	(1) Main	(2) Female	(3) Age	(4) Education	(5) Actual Wage	(6) Experience	(7) Marriage
Wage	0.00108***	-0.000435***	-9.37e-06	-2.93e-06	-2.10e-08*	1.24e-05	-8.18e-05
	(0.000413)	(0.000127)	(7.97e-06)	(2.10e-05)	(1.26e-08)	(1.14e-05)	(0.000172)
Time	-0.0596**	0.0159*	0.00155***	-0.000840	1.52e-07	-0.000263	-0.0127
	(0.0277)	(0.00863)	(0.000519)	(0.00140)	(6.59e-07)	(0.000770)	(0.0113)
Outdoor	0.733	0.395	0.00480	-0.0936	3.38e-06	-0.0220	-0.378
	(1.279)	(0.421)	(0.0260)	(0.0672)	(5.35e-05)	(0.0373)	(0.529)
Nonoffice	-1.406	0.236	0.0584	-0.0393	5.06e-05	-0.104**	0.0328
	(1.741)	(0.534)	(0.0358)	(0.0810)	(6.87e-05)	(0.0506)	(0.610)
Contract	-2.430**	0.148***	0.0378**		1.60e-05	-0.0285	-0.0409
	(1.008)	(0.0492)	(0.0189)		(2.69e-05)	(0.0281)	(0.382)
Insurance	-0.167	0.0615	-0.0224		2.37e-05	0.0269	0.773*
	(1.105)	(0.0560)	(0.0210)		(3.94e-05)	(0.0293)	(0.435)
MidDanger	2.952*	-0.105	-0.0641**		4.91e-05	-0.0313	0.725
	(1.572)	(0.0723)	(0.0304)		(5.57e-05)	(0.0413)	(0.553)
HighDanger	2.902*	-0.630*	-0.0765**	-0.0866	5.07e-05	0.0131	0.170
	(1.608)	(0.376)	(0.0326)	(0.0727)	(6.06e-05)	(0.0424)	(0.579)
FirstLine	-1.951	1.103**	-0.00235	0.0798	6.84e-05	-0.0307	0.234
	(1.647)	(0.461)	(0.0300)	(0.0827)	(5.54e-05)	(0.0475)	(0.646)
SecondLine	0.681	-0.163	-0.0129	0.0388	-1.32e-05	-0.00152	-0.438
	(1.189)	(0.391)	(0.0237)	(0.0600)	(3.96e-05)	(0.0325)	(0.477)

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

We observe from Table 2.4 that compared to male, female migrant workers have weaker preference for wage and are less averse to long working hours. They prefer jobs that provide contracts and dislike jobs that are highly dangerous. Moreover they are more attracted to jobs located in first-line cities. Age plays a very similar role. We see that older migrant workers care less about working hours but prefer more secure jobs – jobs with contracts and no danger. Interaction effects between education and 3 job attributes – contract, insurance, and medium danger – were dropped from the regression due to multicolinearity. However, for migrant workers preference does not seem to vary with education level. Actual wage, experience (years spent in current industry), and marriage have some limited and statistically weak effects on preference patterns. Migrant workers who are currently earning a higher amount tend to care less about wage; more experienced migrant workers tend to dislike non-office jobs; and married people tend to care more about having insurance.

Table 2.5: Interaction Effects (Urban Workers)

VARIABLES	(1) Main	(2) Female	(3) Age	(4) Education	(5) Actual Wage	(6) Experience	(7) Marriage
Wage	-0.00129	-2.15e-05	3.41e-05**	5.70e-05	6.83e-09	-1.87e-05	-0.000317
	(0.00101)	(0.000158)	(1.63e-05)	(4.83e-05)	(2.22e-08)	(1.14e-05)	(0.000290)
Time	0.238***	-0.00416	-0.00339**	-0.0113***	-1.04e-06	0.00163**	0.0352
	(0.0821)	(0.0116)	(0.00133)	(0.00375)	(2.73e-06)	(0.000759)	(0.0236)
Outdoor	-6.966	0.307	-0.0445	0.387**	-0.000123	0.0435	1.628
	(4.261)	(0.608)	(0.0734)	(0.195)	(0.000120)	(0.0438)	(1.019)
Nonoffice	7.271	-0.513	-0.183*	-0.223	-0.000253	0.0615	2.658**
	(6.312)	(0.803)	(0.108)	(0.290)	(0.000257)	(0.0647)	(1.155
Contract	-2.049	0.155	0.0300		-0.000133	0.0190	-0.541
	(2.503)	(0.119)	(0.0387)		(9.68e-05)	(0.0246)	(0.678)
Insurance	4.328	-0.131	-0.00741		9.76e-05	-0.0229	-1.777*
	(3.462)	(0.156)	(0.0593)		(0.000119)	(0.0378)	(0.936)
MidDanger	-6.275	0.297	0.0103		0.000175	-0.0355	1.218
	(4.560)	(0.204)	(0.0720)		(0.000157)	(0.0426)	(1.052)
HighDanger	-9.120*	0.836	-0.0140	0.374	3.76e-05	0.0752	2.113*
	(4.867)	(0.576)	(0.0830)	(0.234)	(0.000192)	(0.0568)	(1.187)
FirstLine	-1.534	-0.294	0.0912	-0.113	5.65e-05	0.0537	-0.754
	(5.043)	(0.708)	(0.0844)	(0.222)	(0.000169)	(0.0454)	(1.348)
SecondLine	-3.169	0.683	-0.0147	0.135	0.000106	0.0477	0.268
	(4.023)	(0.595)	(0.0687)	(0.182)	(0.000153)	(0.0442)	(0.931)

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

Compared to migrant workers, the interaction effects for urban workers are less obvious, both in terms of the number of coefficients that are significant and the strength of statistical significance. Preferences do not vary significantly with gender or current wage level. As for age, older urban workers put more emphasis on wage and show stronger distaste towards overtime and non-office jobs. More educated people tend to care less about working longer hours and display less distaste toward outdoor jobs – which may sound counter intuitive, but is possible since the ones who spent most of their lives sitting in classrooms tend to romanticize outdoor jobs. More experienced urban workers tend to care less about overtime, whereas married workers show less aversion to non-office jobs, jobs without insurance, and highly dangerous jobs.

Estimation in the WTP Space and Results

The Mixed Logit model makes it possible to account for heterogeneity in preferences which are unrelated to observed characteristics and it has been shown that any discrete choice random utility model can be approximated by an appropriately specified mixed logit model (McFadden and Train 2000). Since the WTP for an attribute is given by the ratio of the attribute coefficient to the monetary coefficient, the WTP from a mixed logit model is given by the ratio of two randomly distributed terms. Depending on the choice of distributions for the coefficients this can lead to WTP distributions which are heavily skewed and that may not even have defined moments. A common approach to dealing with this potential problem is to specify the monetary coefficient to be fixed. This is a convenient assumption as in this case the distribution of the willingness to pay for an attribute is simply the distribution of the attribute coefficient scaled by the fixed wage (or price) coefficient. The problem is that this implies that the standard deviation of unobserved utility, which is called the scale parameter, is the same for all individuals. This approach also tends to generate unreasonably large variance of WTPs, which translates into an untenable implication that many people are willing to pay an enormous amount of money to have or avoid an attribute (Train and Weeks, 2005). A more reasonable approach, therefore, is to estimate the Mixed Logit model directly in the willingness to pay space rather than in preference space. This involves estimating the distribution of willingness to pay directly by re-formulating the model in such a way that the coefficients represent the WTP measures. The researcher then makes a priori assumptions about the distributions of WTP rather than the attribute coefficients. This approach has been found to produce more realistic WTP estimates in applications.

The WTP-Space estimation approach, as in Train and Weeks (2005), specifies utility as separable in price (or wage in our case), p, and non-price attributes, x:

$$U_{njt} = -\alpha_n p_{njt} + \beta_n' x_{njt} + \epsilon_{njt}$$

where α_n and β_n are individual specific and ϵ_{njt} is i.i.d. ϵ_{njt} is assumed to be Extreme Value Type 1 with individual specific variance (scale parameter k_n), namely $Var(\epsilon_{njt}) = k_n^2(\pi^2/6)$. Since scale parameter k_n is individual specific, dividing utility by k_n does not affect behavior but will give us a new error term that has the same scale for all observations:

$$U_{njt} = -(\alpha_n/k_n)p_{njt} + (\beta_n/k_n)' x_{njt} + \varepsilon_{njt}$$

where ε_{njt} is iid Extreme Value Type 1 with constant variance $\pi^2/6$. Defind $\lambda_n = (\alpha_n/k_n)$ and $c_n = (\beta_n/k_n)$. Then the preference space estimation equation can be rewritten as:

$$U_{njt} = -\lambda_n p_{njt} + c_n' x_{njt} + \varepsilon_{njt}$$

We can define willingness-to-pay coefficient as the ratio of the attribute's coefficient to the price coefficient, that is $\omega_n = c_n/\lambda_n$ and the WTP-space estimation equation can be written as:

$$U_{njt} = -\lambda_n p_{njt} + (\lambda_n \omega_n)' x_{njt} + \varepsilon_{njt}$$

Under this parameterization, the variation in WTP, which is independent of scale, is distinguished from the variation in the price coefficient, which incorporates scale.

We estimate ω_n directly using this specification. The price parameter $-\lambda_n$ is assumed to follow a log-normal distribution whereas WTP parameters are assumed to be normal. Moreover, the WTP's are assumed to be uncorrelated over attributes.

Table 2.6: WTP Space

	(1)	(2)	(3)
VARIABLES	Migrant	Urban	Difference
Time	-28.90***	-24.18	-2.235
	(9.662)	(19.88)	(17.56))
Outdoor	-1,288**	-5,438***	-3,107***
	(533.4)	(1,791)	(1,157)
Nonoffice	-932.9	-2,245	-570.8
	(706.7)	(1,669)	(1,365)
Contract	1,241***	1,278	-279.5
	(415.6)	(859.1)	(742.5)
Insurance	2,025***	3,196**	-105.8
	(610.4)	(1,500)	(936.4)
MidDanger	-19.09	-520.7	-781.1
	(591.3)	(1,201)	(1,088)
HighDanger	-1,884***	-3,805**	-968.7
	(616.8)	(1,640)	(1,199)
FirstLine	-114.3	284.9	433.0
	(668.7)	(1,423)	(1,311)
SecondLine	1,095**	-675.2	-1,800*
	(529.3)	(1,150)	(1,062)
Observations	1,704	996	2,700

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

Table 2.6 gives the estimation results. All parameters in the table now have monetary interpretations. For example, an average migrant worker needs to be compensated 1,288 RMB to accept an outdoor job while an average urban worker needs to be compensated 5,483 RMB. We observe that all WTP parameters are of reasonable signs although not all of them are significant. Consistent with preference space estimation results, the two statistically significant "difference" estimates are Outdoor and Second Line. The magnitude of the differences, however, is quite large. Urban worker needs to be compensated 3,107 RMB more than migrant workers to accept an outdoor job, and they need to be compensated 1,800 RMB more to be willing to work in a second line city.

In summary, our empirical results substantiate our hypothesis that urban workers are more averse to undesirable job attributes, namely $\beta_U > \beta_R$. Although the difference is not significant for every single attribute, it is large in magnitude when the difference is indeed significant.

2.5 Conclusion

The observation in the Chinese labor market that migrant workers take on blue-collar jobs that urban workers are unwilling to take and are paid less contradicts Rosen's compensating differentials model. It is also not consistent with Lang and Majumdar's more recent prediction that in a labor market with search frictions and heterogeneous preferences, the group that is more averse to unattractive job characteristics will on average earn less. In the theoretical part of this paper, we expand the model by Lang and Majumdar (2003) to show that wages need not be compensating when preferences are heterogeneous, and that the group more averse to undesirable working conditions need not earn less when reservation utilities differ and/or when employers practice taste-based discrimination. In the empirical part, we substantiate the assumption in our theoretical modeling that urban workers are more averse to undesirable working conditions using a discrete choice experiment, where 225 workers in China made hypothetical choices between jobs characterized by different wage levels and working conditions. We backed out preference parameters and willingness to accept measures for job attributes. We find that consistent with our assumption, urban workers need to be compensated more to accept outdoor jobs and jobs in second line cities. We also find that migrant workers have more dispersed preferences that vary with personal characteristics such as gender and education.

This study contributes to the existing literature by, first of all, extending the existing theoretical model and incorporating more realistic features such as heterogeneous reservation utilities and taste-based discrimination either from both sectors or from only one sector. Although in most of the cases the employment or wage outcome predictions are not clear, our model provides multiple possibilities to realign theoretical predictions with empirical observations. We also contribute to the literate by empirically estimating the key parameter in our theoretical model and validating our assumption that urban workers are more averse to undesirable job characteristics than migrant workers. Our choice experiment not only offers a unique dataset but also suggests a robust way to obtain preference parameters and WTA measures for job attributes, which cannot be achieved by using observational data.

One key drawback of this study is that the hypothetical choices that respondents make in our experiments will have no real consequences. This is the typical problem that all stated preference studies have. Although there is no way to eliminate this concern, we reduce its impact as much as possible by offering compensation (20 RMB per respondent) and small gifts to encourage them to make serious considerations. Another potential problem is that our sampling is limited to Shandong Province, which may not be geographically representative of all migrant workers in China.

References:

[1] Adda, J., Dustmann, C., & Stevens, K. (2016). The career costs of children.

[2] Altonji, J. G.,] & Blank, R. M. (1999). Race and gender in the labor market. Handbook of labor economics, 3, 3143-3259.

[3] Benjamin, DJ., Choi, J., & Strickland, A. (2010). Social identity and preferences. The American Economic Review, Vol. 100, No. 4, pp. 1913-1928

[4] Bonhomme, S., & Jolivet, G. (2009). The pervasive absence of compensating differentials. Journal of Applied Econometrics, 24(5), 763-795.

[5] Cheng Z., Smyth R.& Guo F. (2015). The impact of chinas new labour contract law on socioeconomic outcomes for migrant and urban workers. Human Relations, 3, 329352.

[6] Cui F., Ge Y. & Jing F. (2013). The effects of the labor contract law on the chinese labormarket. Journal of Empirical Legal Studies, 3, 462483.

[7] Eriksson, T.,]& Kristensen, N. (2014). Wages or fringes? Some evidence on trade-offs and sorting. Journal of Labor Economics, 32(4), 899-928.

[8] Gallagher M., Giles J., Park A. & Wang M. (2014). Chinas 2008 labor contract law: Implementation and implications for chinas workers. Human Relations, 1, 197235.

[9] Flory, J. A., Leibbrandt, A., & List, J. A. (2014). Do competitive workplaces deter female workers? A large-scale natural field experiment on job-entry decisions. The Review of Economic Studies, rdu030.

[10] Frijters, P., Gregory, R. G., & Meng, X. (2015). The Role of Rural Migrants in the Chinese Urban Economy. Migration: Economic Change, Social Challenge, 33.

[11] Huber, J., & Zwerina, K. (1996). The importance of utility balance in efficient choice designs. Journal of Marketing research, 307-317.

[12] Hwang, H. S., Mortensen, D. T., & Reed, W. R. (1998). Hedonic wages and labor market search. Journal of Labor Economics, 16(4), 815-847.

[13] Kessels, R., Jones, B., & Goos, P. (2011). Bayesian optimal designs for discrete choice experiments with partial profiles. Journal of Choice Modelling, 4(3), 52-74.

[14] Lang, K., & Majumdar, S. (2004). The pricing of job characteristics when markets do not clear: theory and policy implications. International Economic Review, 45(4), 1111-1128.

[15] Li J. (2008). Chinas new labor contract law and protection of workers. Fordham International Law Journal, 3, 10831131.

[16] Li X. & Freeman R. (2015). How does chinas new labor law affect floating workers? British Journal of Industrial Relations, 4, 711735.

[17] List, J. A., & Shogren, J. F. (1998). Calibration of the difference between actual and hypothetical valuations in a field experiment. Journal of Economic Behavior & Organization, 37(2), 193-205.

[18] Luo, J. (2015). Potential and channels in achieving median to high growth in China. Government report, Ministry of Finance, China. Available from URL:

http://www.sem.tsinghua.edu.cn/portalweb/sem?c=fa1u=xyywcn/69292.htm

[19] Mas, A., & Pallais, A. (2016). Valuing Alternative Work Arrangements (No. w22708). National Bureau of Economic Research.

[20] McFadden, D. (1974). The measurement of urban travel demand. Journal of public economics, 3(4), 303-328.

[21] Meng, X. (2000). Labour market reform in China. Cambridge University Press.

[22] Meng, X. (2012). Labor market outcomes and reforms in China. The Journal of Economic Perspectives, 26(4), 75-101.

[23] Meng, X. (2017). The Labor Contract Law, Macro Conditions, Self Selection, and Labor Market Outcomes for Migrants in China. Asian Economic Policy Review, 12(1), 45-65.

[24] Meng, X., & Manning, C. (2010). The great migration in China and Indonesia: trends and institutions. Chapters.

[25] Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. Journal of political economy, 82(1), 34-55.

[26] Rosen, S. (1986). The theory of equalizing differences. Handbook of labor economics, 1, 641-692.

[27] Sandor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers prior beliefs. Journal of Marketing Research, 38(4), 430-444.

[28] Train, K. (2003). Discrete choice methods with simulation. Cambridge

university press.

[29] West, L. A., & Zhao, Y. (Eds.). (2000). Rural labor flows in China (No. 42). University of California Inst of East.

[30] Wiswall, M., & Zafar, B. (2016). Preference for the workplace, human capital, and gender (No. w22173). National Bureau of Economic Research.

[31] Zhao, Y. (2000). Rural-to-urban labor migration in China: the past and the present. Rural labor flows in China, 15, 33.

CHAPTER 3

WHERE DID THE MONEY AND TIME GO? DE-MYSTIFYING THE NEGATIVE IMPACT OF REMITTANCES ON HUMAN CAPITAL INVESTMENT IN THE KYRGYZ REPUBLIC

Abstract

International migration and remittances from overseas may encourage human capital investment and improve educational outcomes in developing countries. Empirical studies, however, have shown mixed evidence some positive, some negative, some have shown zero effects. In our study, we focus on the case of Kyrgyz Republic, one of the largest remittances-receiving countries in Asia. We used a 5-year panel dataset that tracks the same 3,000 households and 8,000 individuals in the country to examine the impact of remittances on both household educational expenditure and attendance rate of school age children. We used instrumental variables and fixed-effects regressions to correct for potential selection bias. We find that remittances have a negative impact on human capital formation namely both educational expenditure and school attendance rate are lower for households that receive a higher amount of remittances. To explore the possible channels of the negative effects, we further regressed itemized household expenditures and the time use pattern of school-age children on remittances. We find that the negative effects can at least be partly attributed to increased expenditure on durable goods and increased hours of child labor on farm work as a compensation for adult labor insufficiency induced by out-migration. Our finding calls for the need for monitoring farm labor hours of school-age children. Implementation and scaling up of financial literacy programs that that help parents balance short-term expenditures (durable goods) and long-term investments (education, health) can be beneficial.

Moreover, targeted investment to improve the quality of education services in the country may help increase perceived return to schooling and may therefore improve human capital investment.

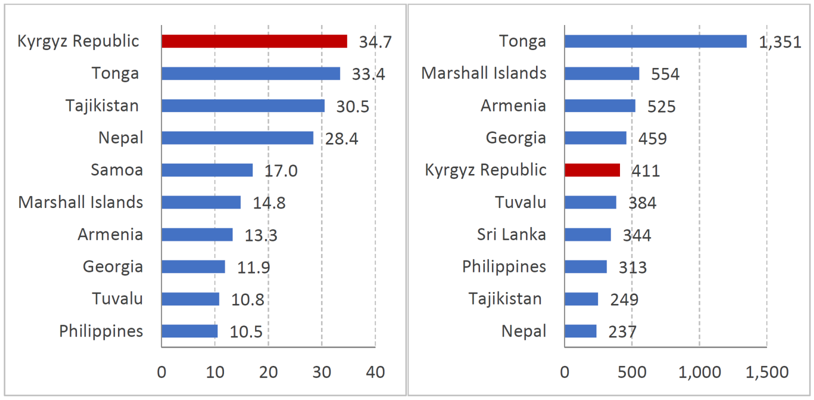
3.1 Introduction

The importance of overseas remittances in developing countries in Asia has been widely acknowledged in the economic literature. Remittances are known to be a major vehicle for economic development. In the short term, it contributes to improve family livelihood and provides immediate disaster relief for the households left behind (Gupta et al, 2009; Acosta et al., 2007; Adams, 2006). In the long term, it may encourage investment in local business and community development. For countries with large young population and low average education level, ascertaining the effects of remittances on human capital investment is especially important. Empirical evidence on such effects, however, has been mixed. Adams et al (2010) and Bansak et al (2009) found strong positive effects in Guatemala and Nepal respectively, while Gregorian et al (2011) and McKenzie and Rapoport (2006) found negative effects in Armenia and Mexico. Other studies, for example Nguyen et al (2015) and Ang et al. (2009) found no significant effects in Vietnam and the Philippines.

A reasonable explanation for the above seemingly contradictory evidence is provided by McKenzie and Rapoport (2006). On the one hand, the inflow of remittances helps to relax households' budget constraint. As a consequence, more resources can be directed to education expenditure. On the other hand, out-migration leads to reduced adult labor force within the household, which may result in increased child labor. The combined effect, therefore, depends on the relative size of the two forces and may vary based on individual and household characteristics (for example, the child's gender, age, number of siblings, and family structure, educational attainments of the parents) as well as country-specific factors (for example, education and labor market systems, legal and institutional frameworks etc.).

In this paper, we estimate the impacts of remittances on human capital investment and educational outcome using a 5-year panel data from the Kyrgyz Republic. The Kyrgyz Republic has very high dependence on remittances. In 2017, its total remittance receipts as a share of gross domestic product was the highest among countries in Asia and the Pacific at 34.7% (Figure 3.1). Per capita remittances during the same year also rank high, sugestive of the substantial potential impacts remittances bring to the country and its households. Moreover, the Kyrgyz Republic has a relatively young population, with its median age at 26.5. Investment in human capital is therefore relevant and crucial for the future development of this country. To date, the empirical evidence is scant as to whether a large inflow of remittances to the country is helping to build human capital with mixed empirical outcomes (Kroeger and Anderson, 2014; WB 2015; Hagedorn, Wang and Chi, 2017), all of which merits revisit on the issue using updated data.

Figure 3.1: Top 10 Remittance-Receiving Economies in Asia – 2017 (% to GDP)



GDP = gross domestic product.
Source: ADB calculations using data from the Global Knowledge Partnership on Migration and Development.

Empirical identification of the effect of remittances on human capital investment is known to be challenging due to the endogenous nature of migration decisions and amount of remittances. To overcome this problem, we first exploit the panel nature of the data to eliminate time-invariant unobserved heterogeneity at the household and individual level. Furthermore, to correct for the possible bias caused by simultaneity, namely the possibility that higher educational expenditure may in turn induce higher remittances, we use a set of instrumental variables (for example: distance to road and other amenities; interaction between land area and incidences of drought) in our panel regressions. We find that remittances have a negative impact on human capital investment and educational achievement in the Kyrgyz Republic even after correcting for endogeneity. Namely, both educational expenditure and school attendance rate are lower for households that receive a higher amount of remittances. To explore the possible channels of the negative effects, we further examined how remittances affected other household expenditures and the time use pattern of school-age children. We find that the negative effects can at least be partly attributed to increased expenditure on durable goods and increased hours of child labor induced by migration and remittances.

Our study contributes to the ongoing research on the impact of remittances/migration on household human capital investment, by offering a rigorous case study of a country with high dependency on remittances but with scarcity of empirical study to date to draw conclusion. The rest of the paper is organized as follows. In Section 3.2, we review previous literature on relevant fields; In Section 3.3, we introduce the dataset we use (Life in Kyrgyzstan, LiK) and provide descriptive statistics to show some general patterns of the data; Section 3.4 presents our identification strategy, instrumental variables used, and estimation results; Section 3.5 discusses the policy implications as well as the limitations of our study and concludes.

3.2 Literature Review

The process by which migration experience and remittance inflows affect human capital investment and educational attainment is complex and multidimensional. It involves a range of social, economic and cultural values and preferences, and is characterized by heterogenous outcomes across jurisdictions, sectors, and individual and household characteristics. Further, remittances and generally the migration process impact the education of children/youth left behind through different channels. Indeed, existing studies show that the net effect of remittances and international migration on human capital investment is ambiguous and heterogeneous at best.

3.2.1 Migration, Remittances, and Educational Investment

According to Gordon H. Hanson and Christopher Woodruff (2003), remittances may have both positive and negative effects on the educational attainment of school-age children, and that there may be differential effects by gender of the children. For example, parents may choose to provide greater access to education for male children, as they may face a greater obligation in providing for elderly parents. yet, boys are a better substitute for lost farm labor due to migration and may be required to forgo education and work more when adults are absent.

A number of empirical studies in the literature support the positive contribution of migration and remittances on education. Remittances relax households liquidity constraints and finance the direct and opportunity costs of schooling, and contribute to household capital accumulation, and higher propensities of migrant families to invest in education (e.g., Acharya and Leon-Gonzalez 2013; Acosta 2011; Adams and Cuecuecha 2010; Bui, Le and Daly 2015; Calero, Bedi and Sparrow 2009; Cox-Edwards and Ureta 2003; Lu and Treiman 2007; Yang 2008). Ultimately, these and other studies (e.g., Bansak and Chezum 2009; Bouoiyour, Miftah and Mouhoud 2016; Bredl 2011; Hanson and Woodruff 2003; Koska et al. 2013; Lopez-Cordova 2006; Mansuri 2006; Salas 2014) conclude that higher income from remittances improve childrens educational outcomes in terms of higher school enrollment, attendance and completion rates; lower dropout rates; and better quality of education.

On the other hand, other studies find reverse and negative effects of the migration and remittance on child education attainment due to the disruptive effects of family migration and other confounding factors. For one, negative effects may be driven by the increased demand for child labor in order to supplement household income or to substitute for the migrant household members domestic responsibilities (e.g., Alcaraz, Chiquiar and Salcedo 2012; Amuedo-Dorantes and Pozo 2010; Antman 2011; Frisancho Robles and Oropesa 2011; Kroeger and Anderson 2014; McKenzie and Rapoport 2011). Moreover, if jobs for migrants are mainly low skilled occupations, left-behind children may discount the value of education, potentially leading to unfavorable educational outcomes. This is further reinforced at the community level if children reside in areas with a high prevalence of international migration. Within these communities, a "culture of migration" can develop, such that young people are expected to migrate in order to attain socioeconomic mobility (Halpern-Manners 2011; Kandel and Massey 2002; McKenzie and Rapoport 2011). Further, some studies point to the emergence of conspicuous consumption among migrant households to the extent of allowing a trade-off between productive and consumptive investment [1]. For instance, based on a review of cross-country case studies, Chami, Fullenkamp and Jahjah (2003) cite that a significant proportion, and often the majority of remittances are spent on "status-oriented" consumption goods. Stephenson and Wilsker (2016) also show that the total effects of remittances are largest in the areas of luxury expenditures and home production and to a lesser extent for education and essential household goods purchases. Additionally, these and other studies (Zhang et al. 2014; World Bank 2015) underline the adverse impacts of the lack or absence of parental inputs, low returns to education both in origin and destination countries, and country-level idiosyncrasies into education investment and acquisition.

[1]Following the logic of Chami et al. (2003), De Brauw and Rozelle (2008), and Yang (2008), among others, consumptive investments include investments in housing and durable goods while productive investments include investments in physical or human capital aimed to increase households income-earning potential (such as in education, agriculture or business).

Overall and notwithstanding the causality direction of the migration and remittance on educational expenditure/outcome discourse, the literature suggests that the magnitude of these effects may vary within specific sociodemographic characteristics and is contingent on/influenced by the interplay of other factors. Some studies find that the potential positive effects of economic migration and remittances on human capital accumulation are much greater for boys, whereas some point to the contrasting pattern favoring girls (Acosta 2011; Bansak and Chezum 2009; Mansuri 2006). Moreover, other studies provide evidence of the interaction of gender and environment differentials, finding positive effects mainly accumulating towards urban males while the negative effects on rural female children (Bucheli, Bohara, and Fontenla 2018). Similarly, some studies note that the impact of the overall migration-remittance process on educational investment is heterogeneously positive, but is most evident among secondary school-aged children and with younger siblings standing to gain the most (Acharya and Leon-Gonzales 2014; Amuedo-Dorantes and Pozo 2010). Among others, this can be attributed to the fact that the direct costs of schooling for primary school-aged children is negligible, given the provision of free public primary schooling.

These mixed results may also reflect a shortcoming found in the identification strategy of existing studies; many of them do not take into account the potential endogeneity and reverse causality. (McKenzie and Yang 2010). This fundamentally reflects the pervasive existence of endogeneity between remittances, migration and developmental outcome variables, of which existing studies either use a variety of instrumental variables or natural experiment to address this endogeneity problem. Across these studies, instrumental variables (IV) technique is the most commonly employed method in addressing the endogeneity of remittances, but the instruments must be carefully selected based on exclusion restriction assumption in order to avoid biased estimates (Gibson and McKenzie and Yang 2013). In this regard, the preferred option has been to rely on natural experiment, where using exogenous economic shocks as a source of variation (Yang 2008) and relying on a visa lottery scheme (McKenzie et al. 2010) have proven to be good examples. Notably, another effective method increasingly employed in the estimation of the impacts of remittances and migration is panel data analysis (Bohme 2015; Funkhouser 2013; Lall et al. 2006; Quisumbing and McNiven 2010; Yang 2008), on which this study is based.

3.2.2 Existing Literature in the Kyrgyz Republic and Other Central Asian Countries

A number of empirical studies in the Kyrgyz Republic have focused on the impact of remittances on overall welfare as well as certain household outcomes. For instance, Karymshakov, Abdieva and Sulaymanova (2014) find that international remittances considerably decrease poverty level in the country. In terms of specific household outcomes, Muktarbek kyzy, Seyitov, and Jenish (2015) analyzed the impact of international migrants remittances on the expenditure structure of households using the Life in Kyrgyzstan survey for 2010-2012, and find that remittances increase the share of expenditures on construction, celebrations, and durable goods, but decrease the shares of expenditures on food and public utilities. Muktarbek kyzy et al.s (2015) findings are consistent with the results of the studies by Ukueva and Becker (2010), who conclude that remittances increase durable goods consumption; and with Hagedorn, Wang and Chi (2017), who find that increases in household remittance received are correlated with spending on a smaller share of the household budget on food and housing, and a larger share on events and other expenses (such as legal and educational expenses).

Studies on the impact of international migration and remittances on human capital investment, particularly on education, have grown over time, but the existing literature in the Central Asia and Kyrgyz Republic is fairly limited. Within Central Asia, Brown et al. (2008) find that school absenteeism increases with the receipt of remittances in Tajikistan, with the negative effect potentially explained by the low level of confidence about the future returns and good employment opportunities. By contrast, Mirkasimov Anderson (2010) find positive effects in terms of overall education expenditure, as well as in terms of school enrolment of older children in the same country while Clement (2011) find no effect. A study by Kroeger and Anderson (2014) on Kyrgyz Republic is the most relevant to our study which evaluates the impact of remittances (domestic and international transfer combined) on child probability of school enrollment using fixed effects and instrumental variable technique. They find, based on data collected between 2005 to 2009 that the receipt of remittances has no significant impact on overall school enrollment, with negative and significant effect on children age 14-18, especially among boys. The negative impact is mainly due to the loss of adult labor in the household. Our study aims to complement this study by making contribution in three areas; 1) isolating the effect of international remittances from that of domestic transfer, 2) providing analysis based on alternative and updated panel data set covering period up to 2016, and 3) introducing additional and more direct outcome variables (educational expenditure and attendance) in addition to the school enrollment data [2], 4) identifying potential channels through which remittances may affect educational expenditure and outcomes, for example expenditure on other goods and changes in childrens time use pattern. Similar to Kroeger and Ander-

[2] WB(2015) explores the impact of international remittances on educational expenditure using OLS regression, but the study does not control for potential endogeneity.

son (2014), we exploit panel structure of the data and IV strategy to identify the causal effects of international remittances on human capital development among school-aged children in the country.

3.3 Data

The data source of our paper is the Life in Kyrgyzstan (LiK) Study, a longitudinal survey that tracks the same 3000 households and 8000 individuals over time in all seven Kyrgyz regions (oblasts, namely Batken, Jalal-Abad, Issyk-Kul, Naryn, Osh, Talas, and Chui) and the two cities of Bishkek and Osh. The data are representative nationally and at the regional level (East, West, North, South). The survey interviews all adult household members about household demographics, assets, expenditure, migration, employment, agricultural markets, shocks, social networks, subjective well-being, and many other topics. The survey was first conducted in the fall of 2010 and has been repeated four times in 2011, 2012, 2013, and 2016. All members of the households in 2010 are interviewed and tracked over time. This implies that if a member of an original sample household leaves the household (e.g. to form an own family), she is still part of the sample. If relevant, other members (e.g. spouse and children) of the new household are then included in the sample as well. By the end of the last wave in 2016, the attrition rate for households is about 20%.

The survey consists of a household questionnaire (to be filled in by the most informed household member), an individual questionnaire (to be filled in by all adults of age 18 and above of the sampled households) and a community questionnaire (to be filled in by a representative of a local administration). Our study utilizes all relevant components of the dataset by merging data across years and modules to create a 5-year unbalanced panel.

Table 3.1 provides some basic summary statistics relevant to the scope of our study. Within the time span of this study, both household income and education expenditure increased steadily over years. With regard to remittances, the percentage of households receiving remittances increased slightly from 2010 to 2013 and subsequently came back to its original level in 2016. Although the percentage of households receiving remittances is only about 10%, which is not high compared to other Asian countries, the receiving households do rely heavily on remittances. As shown in Appendix B, remittances on average accounts for approximately 70% of the annual income of the receiving households. Regardless of measurement used (total vs. per capita, cash vs. in-kind), the amount of remittances received fluctuated greatly over years. This is likely to be due to exchange rate fluctuation and the unstable economic condition of Russia during this period, since the majority of Kyrgyz migrants work in Russia.

Table 3.1: Household Finance Summary Statistics (By Year, in Soms)

	2010	2011	2012	2013	2016
Per Capita Household Income	2,847.01	2,970.17	3,531.47	3,988.62	6,042.15
Percentage Received Remittances	10.00%	12.10%	12.90%	13.90%	10.40%
Total Annual Amount of Cash Remittances	7,587	24,188	22,424	16,533	18,295
Total Annual Amount of Cash Remittances (Receiving Households)	75,867	200,721	175,345	119,507	175,989
Total Annual Amount of In-Kind Remittances	28.67	87.39	124	111.1	83.99
Per Capita Annual Remittances	1,420	4,232	4,429	2,771	2,798
Household Education Expenditure (Primary)	2,044	2,622	2,980	3,781	4,446
Household Education Expenditure (Secondary)	2,826	3,701	4,045	5,249	5,444

Table 3.2 compares the household characteristics of remittances receiving and non-receiving households. As shown in the last column, they are statistically significantly different in all aspects. On the one hand, remittances-receiving households are privileged in the sense that they have higher household income and are more

likely to have male and married household heads. On the other hand, they are disadvantaged since they tend to have less educated household head and are more likely to reside in rural areas. These differences may affect how receiving households perceive the importance of human capital investment and allocate resources between production and education, both in terms of money and in terms of time.

Table 3.2: Remittances Receiving and Non-Receiving Households Comparison

Variables	Non-Receving	Receiving	Difference
Household Size	4.676	6.536	-1.860***
Household Income	18,000	22,000	-4.3e+03***
Male Household Leader	71.20%	76.80%	-0.056***
Household Leader Age	51.67	54.57	-2.896***
Household Leader Education	4.72	4.412	0.308***
Household Leader Married	0.701	0.78	-0.080***
Rural Households	58.10%	74.30%	-0.161***
Observations	12144	1630	

Interestingly, there are huge geographical variations across regions within the Kyrgyz Republic. As shown in Table 3.3, households in the South are more reliant on remittances compared to the North and Central areas. That is, households in the South have more migrants per household, more than 20% of them are receiving remittances, with amounts that are much higher in comparison to other areas.

Table 3.3: Remittance Trend (By Region)

	2010	2011	2012	2013	2016
North (Talas, Baryb, Issyk-Kul)					
Number of Migrants Per Household	0.0476	0.0643	0.0972	0.1200	0.0963
Number of Migrants Per Migrating Household	1.2778	1.3810	1.4074	1.3158	1.3000
Received Remittances	3.43%	4.09%	5.36%	8.02%	6.56%
Total Annual Amount of Cash Remittances	2,691	5,327	5,640	7,134	6,508
Total Annual Amount of In-Kind Remittances	7.62	70.18	99.21	42.19	43.76
Per Capita Annual Remittances	521	1,018	1,316	1,377	1,268
Central (Chui, Bishkek City)					
Number of Migrants Per Household	0.0530	0.0418	0.0428	0.0457	0.0491
Number of Migrants Per Migrating Household	1.3333	1.2333	1.4091	1.4000	1.3200
Received Remittances	2.61%	2.85%	2.14%	1.81%	3.15%
Total Annual Amount of Cash Remittances	2,225	4,792	3,458	1,379	3,225
Total Annual Amount of In-Kind Remittances	7.391	14.25	0	0	0
Per Capita Annual Remittances	495	995	803	273	639
South (Osh, Jalal-Abad, Batken)					
Number of Migrants Per Household	0.3360	0.3570	0.4410	0.4290	0.2880
Number of Migrants Per Migrating Household	1.5238	1.4122	1.6561	1.6304	1.6220
Received Remittances	19.00%	22.80%	24.50%	24.20%	16.40%
Total Annual Amount of Cash Remittances	14,180	47,430	44,214	30,269	31,911
Total Annual Amount of In-Kind Remittances	55.47	153.7	233	211.7	150.7
Per Capita Annual Remittances	2,579	8,138	8,558	4,976	4,692

To a large extent, the school system in the Kyrgyz Republic continues to follow the Soviet model. Education is compulsory for the first 9 years, from approximately age 6/7 to age 15. Following an optional period at a private or state kindergarten, children enroll in primary school for 4 years (age 6/7 to 9). Secondary education is divided into 5 years of Basic Secondary (age 10-15) and Higher Secondary (age 16-17). Primary and Basic Secondary (Grade 1 to 9) are compulsory and are provided at the state institutions free of cost. Upon completing secondary education, a small portion of students would proceed to receive either vocational education or tertiary education. Vocational education is offered through three kinds of courses: A three-year course mixing vocational and general education and preparing for higher education, a two-year course mixing vocational and general education (without preparation to higher education), and a ten-month course of pure vocational education (also open to adults). Vocational education is given in professional lyceum and vocational technical colleges. Tertiary education delivers bachelor degree in four years, which allows students to pursue master programs, lasting two years. PhD programs are offered at some institutes as well. The average enrollment rates for each level of schooling over the survey years are summarized in Table 3.4. We observe that while primary and secondary education has high and stable enrollment rate of approximately 90%, enrollment rate for post-secondary education hovers around 20%. This is likely to be due to the fact that post-secondary education is neither compulsory nor free of charge. In addition to its low level, we also noticed a downward trend in post-secondary enrollment rate over time.

Table 3.4: Enrollment Rate by School Level

	2010	2011	2012	2013	2016	Overall
Primary (Age 6-9)	87.4%	88.4%	84.9%	85.3%	89.5%	87.2%
Secondary (Age 10-17)	93.0%	91.9%	91.6%	92.6%	96.1%	92.9%
Post-Secondary (Age 18-24)	23.4%	21.3%	19.5%	14.7%	17.4%	19.3%

To summarize, although household income and education expenditure have increased over the survey years, school enrollment rate especially post-secondary school enrollment has declined. On the remittances side, we find that although the percentage of households receiving remittances is low, their reliance on remittances is very heavy. The amount of remittances received fluctuated over the survey years and varied greatly across regions. Moreover, we found very distinctive characteristics between remittance-receiving and non-receiving households, which may in turn affect how they perceive the trade-off between education and production. Although the above descriptive statistics provide insights into the basic trends and patterns of our data, we cannot understand the causal relationship without a rigorous identification strategy.

3.4 Estimation Strategy and Results

The impact of remittances on the education of the school-age children in the recipient household is estimated using educational expenditures and school attendance as two outcome variables. Two potential sources of endogeineity may bias our estimate for the impact of remittances on human capital investment and educational outcomes. First of all, time-invariant unobserved characteristics at both household level (neighborhood environment, beliefs about the importance of education, etc) and individual level (ability, personality, etc) may be correlated with both migration/remittances and educational outcomes. We address this problem by using a fixed-effects panel regression methods to difference out these particular factors (Subsection 3.4.1). Secondly, time-variant unobserved factors and reverse causality – the possibility that higher school costs may induce family members to remit more – cannot be addressed by fixed effects alone. We therefore use an instrumental variable approach to achieve clean identification (Subsection 3.4.2).

3.4.1 Fixed-Effects Panel Regression

The amount of remittances received (or even the decision to migrate) and human capital investment decisions may be simultaneously affected by unobserved household characteristics, such as the members latent ability, beliefs about the importance of education, and neighborhood characteristics. Similarly, a childs unobserved ability and personality may be correlated with both her parents migration decisions and the probability of her attending school. Such endogeneity, left uncorrected, can lead to biased estimate of the effect of remittances on human capital investments. We therefore exploit the panel feature of our data and use fixed effects model to eliminate time-invariant unobserved household characteristics. More specifically, our education expenditure regression model can be written as:

$$ln(y_{jt}) = \alpha_0 + \alpha_j + \beta \cdot Remmit_{jt} + \gamma_t \cdot T + \delta \cdot X_{jt} + \varepsilon_{jt}$$

where the dependent variable $ln(y_{jt})$ is the natural log of the educational expenditure (up to secondary education only) of household j in year t. The key independent variable $Remmit_{jt}$ – the amount of remittances received by household j in year t – takes two forms in our specification: the natural log of total remittances received by the household, and the natural log of per capita remittances. The logarithm transformation is common in econometric studies using wage or remittances data since the residuals have a strongly positively skewed distribution. Furthermore, the log-log specification gives the coefficient an elasticity interpretation. In the above equation, α_j is an $j \times 1$ vector capturing household-specific time-invariant fixed effects, whereas T captures year fixed effects. X_{jt} is a matrix of control variables including characteristics of the household head (gender, age, education level, marital status, and ethnicity), household size, and house value as a proxy for wealth. Some of the control variables are almost time-invariant, such as gender and ethnicity of the household head and only see changes when the household head in previous years leaves the sample or a new household is formed. The variations that we can exploit from these variables are therefore very limited.

For our educational outcome specification, the likelihood of a child attending school is a function of household remittances, a vector of individual child characteristics (the child's gender and age), and household characteristics (gender, age, education level, marital status, and ethnicity of the household head, household size, and house value). Namely,

$$S_{ijt} = \alpha_0 + \alpha_i + \beta \cdot Remmit_{jt} + \gamma_t \cdot T + \delta \cdot C_{it} + \delta \cdot X_{jt} + \varepsilon_{ijt}$$

where S_{ijt} is a binary indicator that takes value 1 if child i (age 5-24) in household j is attending school in time period t. Same as the human capital investment specification, $Remmit_{jt}$ is the amount of remittances received by household j in year t and is included either as the logarithm transformation of total remittances or per capita remittances. To save space, only the results from per capita remittances specifications are reported in the regression tables. C_{it} is the child's characteristics, and α_i is an $i \times 1$ vector capturing child-specific time-invariant fixed effects. X_{jt} and T are the same as defined in the human capital investment specification. Descriptive data of the variables used in estimation are found in the Appendix.

3.4.2 Choice of Instrumental Variables

Although the fixed effects model partials out the effects of time-invariant unobservables and alleviates omitted variable bias, the endogeneity problem is not fully addressed for the following two reasons. First of all, we cannot preclude the existence of time-variant omitted variables affecting both remittances received and human capital investment/ educational outcomes. Secondly, the need to invest more in education may reversely trigger an increase in remittances, thereby biasing the coefficient estimate. To address these problems, we use a set of instrumental variables (IVs) to correct for potential endogeneity bias.

Existing literature has identified variables such as distance to railroad lines, historical migration rates, exogenous shocks to agricultural production such as changes in rainfall patterns (e.g.,Woodruff & Zenteno, 2007, Hanson & Woodruff, 2002, McKenzie & Rapoport, 2007, Munshi, 2003 and Passel, 2006) as valid instrumental variables. We focus on these variables and modify the specifications according to the context of our regressions.

Three sets of instrumental variables were used in our analysis, one at a time depending on the context of the regressions. The first set is land area interacted with severity of drought (no drought=0; mild drought=0.5; severe drought =1), where the incidence and severity of drought is self-reported by each household. We drew inspiration from Munshi (2003) where he used rainfall in the origin-community (collected from local weather stations) as an instrument for the size of the migrant network at the destination, because low rainfall at the origin increases the probability that the migrant will be occupied in a non-agricultural job. Yang and Choi (2007) also used rainfall shocks to instrument for income and remittances in the Philippines. In our case, the drought may lead to increased out-migration due to reduced farm jobs and may lead to higher remittances to compensate reduced income from agricultural production. By interacting land area and severity of drought, we aim to measure the impact of the disaster more accurately, since larger farms suffer heavier loss from the drought. One criticism for the use of natural disasters as instrumental variable in remittance studies is that while drought or rainfall are exogenous shocks that affect remittances, it may also affect household expenditures. While we acknowledge the possibility that other household expenditures may be affected, educational expenditure, which is paid by semester, is unlikely to vary with natural disasters. Hence we only use this set of instrumental variable when the dependent variable is educational expenditure.

The second set of instrumental variable is household's self-reported distance to the nearest road. This type of instrument (distance to the nearest railroad) has been used before in the literature by Woodruff and Zenteno (2007) for the case of Mexico and by Adams and Cuecuecha (2010) for the case of Guatemala. The rationale is that distance to the nearest transportation system is a good proxy for migration costs. Since road is the major mode of transportation in Kyrgyz Republic, the further away a household is from the road system, the lower the likelihood any household member would migrate and hence lower remittances. One potential drawback is that distance to the nearest road tends to be time-invariant. In our dataset, however, we do have enough within-subject variation to allow for identification. The change may either be induced by construction of new roads or households moving to new locations. This set of instrument is used for the regression of other household expenditures on remittances, since distance to the nearest road is unlikely to be correlated to expenditures on durable goods, food, wedding, etc.

For the estimation of the impact of remittances on school attendance, distance to the nearest road may not be a valid instrument as it is likely to directly affect the transportation cost of childrens commute to school, and is therefore likely to be correlated with both educational expenditure and school attendance rate. We therefore introduce a third set of instrumentsthe average distance from the household to the nearest road, market, town hall, and hospital. While distance to the nearest road proxies for migration costs, the averaged distance measure proxies for migration costs and the general centrality of the households location and, at the same time, reduces the correlation with school attendance.

The first-stage regression results are included in Appendix B. For each regression, the instrumental variable included is statistically significant ($p < 0.001$) on its own, Moreover, all the three first stage regressions are significant ($p < 0.10$) and have passed the weak IV tests. In the results section for our main regressions, we report both the fixed-effect panel regression results without correcting for endogeneity (columns labeled FE) and with endogeneity corrected (columns labeled"FE, IV) to allow for side-by-side comparison. For auxiliary regressions that aims to identify the channels of the impact of remittances, we only report the results after correcting for selection bias.

3.4.3 Estimation Results (1): Remittances and Human Capital Investment

The estimation results for the effect of remittances on human capital investment, measured by household level educational expenditure, are summarized in Table 3.5. The fixed effects panel regression results without using instrumental variables are presented in columns (1) and (3), whereas the endogeneity corrected results are presented in columns (2) and (4). The first two columns use the natural log of total remittances received by the household as the independent variable, while the last two columns use the natural log of per capita remittances received by the household.

We observe that remittances have a negative effect on household educational expenditure. That is, households that receive more remittances tend to spend less on the education of children. Although the effect is not robust to endogeneity correction as the key estimates in columns (2) and (4) are not statistically significant, the trend is negative and very close to significant. In addition, educational expenditure is also affected by household characteristics. For example, older household heads tend to spend more on childrens education, and wealthier households (proxied by market value of the house) have higher educational expenditure.

Table 3.5: Effect of Remittances on Education Expenditure

VARIABLES	(1) HHD Educ Exp FE	(2) HHD Educ Exp IV,FE	(3) HHD Educ Exp FE	(4) HHD Educ Exp IV, FE
Remittances				
Log Total Remittances	-0.0218*	-0.289		
	(0.0131)	(0.272)		
Log Per Capita Remittances			-0.0262*	-0.349
			(0.0157)	(0.327)
Household Characteristics				
Age of Household Leader	0.0289***	0.0246*	0.0288***	0.0235*
	(0.0103)	(0.0141)	(0.0103)	(0.0141)
Gender of Household Leader	-0.580	-1.401*	-0.581	-1.403*
	(0.399)	(0.833)	(0.399)	(0.835)
Household Size	-0.0184	0.108	-0.0188	0.103
	(0.0507)	(0.0702)	(0.0507)	(0.0694)
Education Level of Household Leader	0.0254	0.0484	0.0257	0.0519
	(0.0912)	(0.110)	(0.0912)	(0.109)
Marital Status of Household Leader	0.258	0.649	0.259	0.650
	(0.320)	(0.458)	(0.320)	(0.459)
House Value	2.37e-07***	2.17e-07**	2.38e-07***	2.19e-07**
	(4.31e-08)	(8.83e-08)	(4.31e-08)	(8.98e-08)
Ethnicity Control	Yes	Yes	Yes	Yes
Constant	6.694***	6.926***	6.696***	6.996***
	(0.805)	(1.321)	(0.805)	(1.363)
Observations	4,828	3,925	4,828	3,925
Number of HHD	1,980	1,781	1,980	1,781

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

The negative (or lack of) impact of remittances on educational expenditure seems counterintuitive. One possible explanation is that remittances might have induced an increase in expenditure on certain items that requires existing household resources to be drawn into the new items. For example, the household may decide to buy a new car after receiving remittances. However, the amount of remittances received may not be sufficient for the car, so the household head will have to cut expenditure on schooling or healthcare to fund the new car. To understand how household expenditure on other items changed in response to remittances, we performed a set of regressions of itemized expenditures on remittances, using distance to the nearest road as instrumental variable . The categorization method of household expenditure items is summarized in Appendix B. Table 3.6 reports the endogeneity corrected results.

Table 3.6: Effect of Remittances on Other Household Expenditure (FE, IV)

VARIABLES	(1) Food	(2) Non-Durable	(3) Durable	(4) Wedding	(5) Utilities	(6) Health	(7) Other
Remittances							
Log Total Remittances	-0.0116	-4.018***	4.605***	-0.194	0.00757	1.085***	0.382**
	(0.0559)	(1.147)	(1.290)	(0.344)	(0.0936)	(0.367)	(0.174)
Household Characteristics							
Age of Household Leader	0.0103***	0.0722**	0.0963***	-0.0355***	0.00806***	0.0375***	-0.00302
	(0.00157)	(0.0320)	(0.0360)	(0.00960)	(0.00261)	(0.0102)	(0.00484)
Gender of Household Leader	-0.0515	-1.135	2.163*	-0.265	-0.227**	0.320	0.178
	(0.0556)	(1.130)	(1.271)	(0.339)	(0.0922)	(0.362)	(0.171)
Household Size	0.0971***	0.653***	-0.225	0.151**	0.0651***	0.00661	0.114***
	(0.00994)	(0.197)	(0.222)	(0.0591)	(0.0161)	(0.0631)	(0.0298)
Education Level of Household Leader	0.0208	-0.510*	-0.250	0.0755	0.00457	0.00649	-0.0484
	(0.0138)	(0.282)	(0.317)	(0.0846)	(0.0230)	(0.0903)	(0.0427)
Marital Status of Household Leader	0.0997**	0.421	-1.055	-0.293	0.363***	0.231	0.0418
	(0.0418)	(0.854)	(0.959)	(0.256)	(0.0696)	(0.273)	(0.129)
House Value	7.99e-08***	7.63e-07***	1.42e-06***	6.51e-08	3.18e-08***	2.08e-07***	9.91e-08***
	(6.63e-09)	(1.35e-07)	(1.52e-07)	(4.05e-08)	(1.10e-08)	(4.32e-08)	(2.04e-08)
Ethnicity Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	9.896***	8.218***	-4.078	7.568***	8.669***	3.329***	8.417***
	(0.128)	(2.616)	(2.940)	(0.784)	(0.213)	(0.837)	(0.396)
Observations	12,824	12,871	12,871	12,871	12,871	12,871	12,871
Number of HHD	3,090	3,093	3,093	3,093	3,093	3,093	3,093

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

We find that remittances have a negative impact on non-durable goods (e.g., clothing, shoes, personal care items) but have positive impacts on durable goods (e.g., cars, phones, computers, TV); healthcare (e.g., medicine, hospital visits); and other expenditures (e.g., house maintenance, recreation, taxes). Interestingly, expenditure on durable goods has the highest elasticity - a 1% increase in remittances may lead to an approximately 4% increase in durable goods expenditure.

This suggests that the negative response of educational expenditure to remittances may be induced by the large expansion in demand for durable goods. This also means that an increase in remittances is less likely to lead to a parallel increase in expenditures for education and non-durable goods because of the greater demand to purchase durable consumer goods. This is consistent with the findings of World Bank (2015) which reports that while being a migrant household correlates with higher consumption, it does not correlate with higher education expenditures (either in total or in per capita terms using OLS estimation controlling for household characteristics). In addition, noting that the dependency ratio for migrant households is lower than that of non-migrant households, there is evidence that the choice of migration is not necessarily investment into youth but rather on asset accumulation (ibid). Meanwhile, Muktarbek kyzy et al. (2015) report partially similar results, noting that remittances have a greater effect on expenditures on durable goods than on education (i.e., a 1% increase in remittances share from total income increases the share of expenditure on durable goods by 1.12% and on human capital by 0.7%).

In summary, our findings suggest that migrant households in the Kyrgyz Republic are primarily driven by consumptive investment (i.e., investment in assets and goods that immediately improve the quality of life and standard of living of households) rather than productive investment (i.e., investment in assets that improve the productive capacity of households in the long-run). This may be partially explained by the fact that migrant households in the country have low levels of domestic sources of income and social assistance, are increasingly becoming dependent on remittance income, and thus are more vulnerable to economic shocks (World Bank 2015). In this regard, remittance income is then more likely to be used for immediate consumption or as shock absorber as substitute for social protection when the domestic labor market or social assistance is poor (Kireyev 2006). However, the pattern of expenditure on durable goods and other items such as recreation and house maintenancewhich are often associated as status-oriented consumption goodsmay suggest that the receipt of remittances may, at the same time, allow conspicuous consumption at the expense of long-term productive investments.

On another note, within household characteristics, the finding that older household heads tend to have a larger expenditure allocation on education seems to highlight the factor of maturity at play in influencing household expenditure and consumption behaviors, more so on the productive activities of families. In addition, the finding relating to investment differences by wealth levels indicate that wealthier households may have better access to credit, which allows them to engage in more investment opportunities, including on education. Notwithstanding this possibility, it is worthy to note that overall, expected returns to education and country-level idiosyncrasies tend to play a large role in determining households education expenditures in the Kyrgyz Republic, especially towards post-secondary education. A recent study by Esenaliev (2018) cites that an increasing share of secondary school graduates opt to leave schools, in part, due to higher admission requirements in higher education institutions and lack of appropriate employment prospects for university graduates. In addition, despite the increase in higher education institutions in recent years, the quality of higher education remains poor. Even while a university diploma does provide better chances of being employed, unemployment among those with higher education remains considerably high, at about 18% in 2015. This contextual evidence that the quality of and returns to education is among, if not the major, determinant in Kyrgyz households expenditure allocation on education is partially consistent with the findings of Clement

(2011) and Brown et al. (2008).

3.4.4 Estimation Results (2): Remittances and Educational Outcomes

While educational expenditure is a good measure on households investment in human capital, educational outcomes are of utmost importance. In this section, we evaluate the effect of remittances on educational outcomes, in particular, school attendance rate of school-age children/youth. The dependent variable is a binary variable that takes the value of 1 if a person aged between 5 and 24an age range that covers primary schooling to Masters educationis currently attending school and 0 otherwise. Since the dependent variable is at the individual level, we use per capita remittances received instead of total remittances received as our key independent variable.

Table 3.7: Effect of Per Capita Remittances on Attendance Rate (By School Level)

VARIABLES	(1) Overall FE	(2) Overall IV, FE	(3) Primary FE	(4) Primary IV, FE	(5) Secondary FE	(6) Secondary IV, FE	(7) Post FE	(8) Post IV, FE
Remittances								
Log PC Remittances	-0.0531***	-1.136***	0.0303	0.263	-0.0290*	-0.605***	-0.0411**	-0.0507
	(0.00883)	(0.147)	(0.0346)	(0.492)	(0.0165)	(0.211)	(0.0199)	(0.406)
Child Characteristics								
Age	-0.332***	-0.246***	-0.595***	-0.619***	-0.0760***	-0.0291	-0.626***	-0.625***
	(0.0130)	(0.0171)	(0.0867)	(0.0970)	(0.0255)	(0.0306)	(0.0441)	(0.0577)
Gender	0.0361	-0.333	1.798	1.716	-0.525	-0.712	-0.100	-0.0786
	(0.254)	(0.261)	(1.354)	(1.329)	(0.461)	(0.470)	(0.647)	(0.670)
Household Characteristics								
Household Size	-0.0246	0.0456	-0.152	-0.174	0.0957*	0.133**	-0.0162	-0.0238
	(0.0292)	(0.0311)	(0.122)	(0.127)	(0.0574)	(0.0593)	(0.0673)	(0.0731)
Age of Household Leader	0.00716	0.0249***	0.00342	0.000423	0.00739	0.0171	-0.0194	-0.0193
	(0.00619)	(0.00664)	(0.0270)	(0.0288)	(0.0115)	(0.0119)	(0.0148)	(0.0164)
Gender of Household Leader	0.252	-0.337*	0.139	0.138	-0.248	-0.563	0.158	0.134
	(0.179)	(0.197)	(0.958)	(0.965)	(0.335)	(0.353)	(0.482)	(0.537)
Education Level of Household Leader	0.0262	0.0272	-0.278	-0.251	0.167*	0.181*	-0.124	-0.125
	(0.0516)	(0.0518)	(0.220)	(0.218)	(0.101)	(0.101)	(0.105)	(0.104)
House Value	-4.02e-07***	-4.21e-07***	-1.16e-06***	-1.15e-06***	-6.07e-07***	-6.23e-07***	-3.41e-07***	-3.44e-07***
	(3.27e-08)	(3.29e-08)	(1.71e-07)	(1.70e-07)	(6.03e-08)	(6.09e-08)	(7.66e-08)	(7.73e-08)
Ethnicity Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	11,012	11,011	1,017	1,017	3,376	3,376	2,301	2,301
Number of Individuals	2,724	2,724	393	393	948	948	683	683

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

The first two columns of Table 3.7 summarize our estimation results using the full sample, with and without instrumental variables. We observe that the coefficient on remittances remains negative and statistically significant regardless of the model used. That is, remittances have a negative impact on childrens school attendance rate.

We further breakdown the effects by school level and report them in columns (3) to (8) in Table 3.7. We observe that regardless of the model used, remittances do not have a statistically significant impact on primary school attendance rate (columns (3) and (4)). This is not surprising as primary education is compulsory and children between ages 6 and 9 are unable to substitute for adult labor. In terms of secondary school attendance rate, the effect is, however, negative and statistically significant. This may be associated to the maturity of children aged

10-17, i.e., their ability to act as potential substitutes for adult labor. To note, in the LiK questionnaire, household heads are asked to report if, and why, school-age children within their households are not attending school. While works to support family is not considered as a reason for any absence from primary school, it is quoted as a reason for approximately 10% of the cases of secondary school non-attendance. For post-secondary education, the effect seems to be negative but is not statistically significant after correcting for selection bias. One possible explanation rests on the fact that college attendance rate in the Kyrgyz Republic is generally very low [3]. Therefore, the effects of remittances on such small variation can hardly be robustly identified.

As mentioned in the previous section, the attractiveness of higher secondary (non-compulsory) and post-secondary education in the country is on a downward trend due to higher admission requirements and low quality and returns in terms of prospective employment opportunities. Similarly, the poor rewards to education in destination countries may have a disincentive effect on school attendance, and ultimately on educational attainment. Kyrgyz labor migrants are generally less educated than domestic workers [4], and most are engaged in low-skilled occupations (World Bank 2015). A decline in migration is also apparent among households in the higher wealth quantiles while poorer households are increasingly becoming migrant-sending households (ibid). These two patterns do not only underline the fact that labor migration in the Kyrgyz Republic exhibits a selection bias towards younger, less educated workers, but also that returns to education in destination countries are generally low.

[3] In fact, enrollment in tertiary education has been flat since 2009 (World Bank 2015).

[4] About two-thirds of labor migrants have completed a general secondary degree, while over a third of domestic workers have completed a specialized secondary or tertiary degree (World Bank 2015).

On another note, school attendance rate is also negatively affected by age, with older children less likely to attend school. This is in parallel with World Banks (2015) study which reports that older children aged 1518 are less likely to be enrolled in school. Gender surprisingly does not seem to play a role in determining school attendance. Interestingly, house value also negatively affects attendance rate, indicating that children from wealthier families are less likely to attend school. This might be due to the possibility that households that own larger properties require kids to help more with house work or farm work. Since house value is not exogenous, it might also be that some common unobservables are driving both wealth and attendance rate.

An alternative breakdown method is by compulsoriness, where we divide the school-age children sample into compulsory education (age 6-15) and non-compulsory education (age 16 and above). The results are reported in Table 3.8. We observe that for both categories -compulsory and non-compulsory - and across specifications used, the impact of remittances on school attendance rate is uniformly negative and statistically significant. In other words, the existence of compulsory education does not negate the negative impact of migration and remittances on childrens educational outcomes. It is understandable that the effects are statistically significant for both categories, since secondary education - the school level that is most heavily impacted by migration and remittances - is included in both categories. In terms of individual characteristics and household characteristics, the finding is the same as in the by school level breakdown analyses. We find that older children and children from wealthier families are less likely to attend school, both in compulsory and non-compulsory education.

Table 3.8: Effect of Per Capita Remittances on Attendance Rate (By Compulsoriness)

VARIABLES	(1) Compulsory FE	(2) Compulsory FE, IV	(3) Non-Compulsory FE	(4) Non-Compulsory FE, IV
Remittances				
Log Per Capita Remittances	-0.0268*	-0.918***	-0.0681***	-0.791***
	(0.0139)	(0.186)	(0.0161)	(0.277)
Child Characteristics				
Age	-0.0348*	0.0400	-0.842***	-0.774***
	(0.0205)	(0.0255)	(0.0350)	(0.0422)
Gender	-0.217	-0.525	-0.142	-0.422
	(0.363)	(0.370)	(0.576)	(0.585)
Household Characteristics				
Household Size	-0.00321	0.0644	-0.0506	-0.0161
	(0.0458)	(0.0478)	(0.0578)	(0.0611)
Age of Household Leader	0.00769	0.0217**	-0.00324	0.00885
	(0.00872)	(0.00920)	(0.0130)	(0.0137)
Gender of Household Leader	-0.106	-0.551*	0.402	-0.00859
	(0.285)	(0.303)	(0.382)	(0.409)
Education Level of Household Leader	0.0660	0.0717	-0.0337	-0.0285
	(0.0765)	(0.0772)	(0.0885)	(0.0877)
House Value	-5.65e-07***	-5.89e-07***	-3.03e-07***	-3.26e-07***
	(5.01e-08)	(5.08e-08)	(6.44e-08)	(6.43e-08)
Ethnicity Control	Yes	Yes	Yes	Yes
Observations	4,668	4,668	4,556	4,555
Number of Individuals	1,264	1,264	1,259	1,259

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

The negative effect of remittances on school attendance rate is similarly counterintuitive as in the educational expenditure case. It is therefore an empirical question why school-age children in high remittances families are less likely to attend school and, if they are not in school, where was their supposed time for education/schooling spent. To answer these questions, we performed an additional set of analyses where we aim to identify the effect of remittances on school-age childrens time use pattern.

In the household questionnaire, household heads are required to answer the following three questions: (1) On average, how many hours each day did [CHILD NAME] spend doing homework during the past academic year (we term it homework); (2) On average, how many hours each day did [CHILD NAME] spend helping at home, family business or farm during the past academic year (we term it housework); and (3) If any, how many hours a day did [CHILD NAME] work outside of the household for money in the past academic year (we term it outside work). We regress these three measures on per capita remittances using the fixed effects model with and without instrumental variables. The instrument used is the average distance from the household to road, town hall, hospital, market, etc.the reason being that the average distance is likely to affect migration cost, and hence remittances, but is unlikely to affect the amount of time that children spend on homework, housework, and outside work. It must be noted that only children between ages 5 and 17 are included in this analysis since time use patterns for children beyond 17 years old are unreported. The estimation results are presented in Table 3.9, with the plain fixed effects model (FE) and instrumented fixed effects model compared in juxtaposition to each other.

Table 3.9: Effect of Per Capita Remittances on Children's Time Use Pattern

VARIABLES	(1) Homework FE	(2) Homework FE,IV	(3) Farmwork FE	(4) Farmwork FE,IV	(5) Outside Work FE	(6) Outside Work FE,IV
Remittances						
Log Per Capita Remittances	-0.000349	0.00872	0.0153***	0.154*	-0.00109	0.0157
	(0.00421)	(0.0596)	(0.00551)	(0.0809)	(0.00226)	(0.0320)
Household Characteristics						
Age	0.0999***	0.0996***	0.173***	0.164***	0.0240***	0.0230***
	(0.00604)	(0.00706)	(0.00790)	(0.00959)	(0.00324)	(0.00380)
Gender	0.132	0.143	0.201	0.254	-0.00163	0.00436
	(0.113)	(0.115)	(0.148)	(0.156)	(0.0606)	(0.0618)
Household Size	0.00596	0.00433	0.0338*	0.0100	-0.0269***	-0.0298***
	(0.0144)	(0.0177)	(0.0188)	(0.0240)	(0.00771)	(0.00950)
Age of Household Leader	0.00264	0.00257	-0.00108	-0.00246	0.00191	0.00175
	(0.00263)	(0.00270)	(0.00344)	(0.00366)	(0.00141)	(0.00145)
Gender of Household Leader	-0.164**	-0.158*	-0.0766	0.0104	-0.0552	-0.0447
	(0.0809)	(0.0891)	(0.106)	(0.121)	(0.0433)	(0.0479)
Education Level of Household Leader	0.0373	0.0376	0.0261	0.0296	0.00550	0.00593
	(0.0229)	(0.0229)	(0.0299)	(0.0311)	(0.0123)	(0.0123)
House Value	6.18e-09	6.17e-09	3.28e-08*	3.66e-08**	-1.84e-08***	-1.79e-08**
	(1.31e-08)	(1.32e-08)	(1.71e-08)	(1.79e-08)	(7.00e-09)	(7.09e-09)
Ethnicity Control	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.330	0.326	-1.074***	-1.028***	-0.0986	-0.0932
	(0.216)	(0.218)	(0.282)	(0.296)	(0.116)	(0.117)
Observations	12,413	12,408	12,413	12,408	12,413	12,408
Number of ID	4,608	4,608	4,608	4,608	4,608	4,608

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

From the first row, we observe that while the coefficients on homework and outside work are not statistically significant, the coefficients on housework are significantly positive. This implies that school-age children in households that receive higher remittances spend a significantly higher amount of time helping at home, family business or farm. This result is robust regardless of which model is used. The magnitude of the impact looks rather small at first sighta 10% difference in annual remittances translates into 30 minutes per month, or equivalently 6 hours per year, difference in a given school-age childs time spent on housework. However, this also implies that if household A receives twice as much remittance as household

B does, the child in household A spend 60 hours more on housework annually than the child in household B, which is a significant amount of time.

Our result is consistent with McKenzie and Rapoport (2006) in that while remittances may help relax households' budget constraint, the absence of adult labor force calls for child labor as a substitute, alters school-age childrens time use pattern, and ultimately negatively affects school attendance rate. In the Kyrgyz Republic, the labor substitution effect outweighs the budget constraint relaxation effect, as households tend to spend the extra income into durable-goods purchases rather than on human capital investments. It is interesting to note that the labor substitution effect only occurs in housework, not outside work. One major reason may be the tightening of regulations on child labor in recent years. In 2004, the Kyrgyz Republic ratified the International Labor Organization Convention for the Elimination of the Worst Forms of Child Labor. In 2011, the Parliament strengthened the Criminal Code by increasing penalties for adults found guilty of crimes against children, including enslavement. In addition, the Government adopted the 2012 - 2014 Social Protection Development Strategy and Action Plan, which serves to protect children and families in difficult conditions, including child laborers. Such laws and regulations increase the risk and cost of hiring children on the formal labor market. Housework, the blind spot of governmental monitoring, is thus a rather convenient channel of substituting adult labor with child work.

In addition to the remittances panel, we observe from the household characteristics panel that older children spend more time on all three types of activitieshomework, housework, and outside workmost likely because they have more school work to do and, at the same time, better ability to perform household duties and jobs outside the household. Children in larger households spend less time on outside work, potentially because adult labor supply is sufficient with more family members working. Children in households headed by males tend to spend less time on homework, which is an interesting finding with little theoretical foundation. It is likely that female household heads put greater emphasis or have favorable preferences toward supporting childrens schooling/education. It is also likely that male household heads tend to spend more time for work and have thus less time for parental monitoring for childrens school performance. Lastly, we find that while children in wealthier households spend less time doing outside work, they do spend more time on housework. Moreover, the magnitude of the effect of wealth (proxied by house value) on housework outweighs that on outside work, meaning that children from wealthier households tend to work more overall. This helps explain the seemingly puzzling negative effect of wealth on school attendance rate school age children in wealthier families are working instead of attending schools. The underlying reason is theoretically ambiguous, but it is likely that parents in wealthier families spend more time earning wages, leaving school-age children at home to attend to housework and substitute for parents domestic responsibilities.

Inspired by Hanson and Woodruffs (2003) finding that remittances may have different effects on boys and girls, we are interested in testing if the effects of remittances on school-age childrens educational outcomes vary by gender. As explained in the literature review section, the theoretical prediction is inconclusive. On the one hand, parents may prioritize the education for boys, since they have better employment opportunities on the labor market. On the other hand, boys are a better substitute for lost farm labor due to migration. To empirically determine the relative magnitude of the above two effects, we perform two additional sets of regressions. First, we examine if remittances have differential impacts for boys and girls on school attendance rate. Secondly, we examine if the impact of remittances on children's time use pattern differ by gender.

In Table 3.10, we separate the sample by gender and look at the effects of remittances on boys and girls. The coefficients are uniformly negative and statistically significant, regardless of models used. However, we observe that the magnitude of the coefficients for boys (columns (1) and (2)) is larger in absolute value than for girls ((columns (3) and (4)), and the relationship is true for both plain fixed effects specification and the instrumental variable specification. This implies that although school attendance rates for both boys and girls are negatively impacted by migration and remittances, the detrimental effects are greater for boys compared to girls. This empirically shows that in the case of the Kyrgyz Republic, labor substitutability outweighs the prospect of future employment, which results in boys being burdened more with housework.

Table 3.10: Effect of Remittances on Attendance Rate (By Gender)

VARIABLES	(1) Attendance Boys FE	(2) Attendance Boys FE, IV	(3) Attendance Girls FE	(4) Attendance Girls FE, IV
Remittances				
Log Per Capita Remittances	-0.0719***	-1.194***	-0.0289**	-1.067***
	(0.0124)	(0.208)	(0.0130)	(0.208)
Individual Characteristics				
Age	-0.327***	-0.237***	-0.348***	-0.267***
	(0.0181)	(0.0237)	(0.0195)	(0.0251)
Household Characteristics				
Household Size	0.000424	0.0729	-0.0644	0.00601
	(0.0418)	(0.0448)	(0.0424)	(0.0448)
Age of Household Leader	0.00350	0.0234**	0.0118	0.0278***
	(0.00858)	(0.00926)	(0.00923)	(0.00980)
Gender of Household Leader	0.232	-0.383	0.226	-0.338
	(0.250)	(0.275)	(0.266)	(0.291)
Education Level of Household Leader	0.0340	0.0498	0.0185	0.0100
	(0.0755)	(0.0755)	(0.0719)	(0.0724)
House Value	-4.19e-07***	-4.37e-07***	-3.92e-07***	-4.14e-07***
	(4.60e-08)	(4.59e-08)	(4.79e-08)	(4.83e-08)
Ethnicity Control	Yes	Yes	Yes	Yes
Observations	5,719	5,718	5,143	5,143
Number of Individuals	1,399	1,399	1,301	1,301

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

We are therefore interested in investigating if the larger negative effect of remittances on boys' attendance rate is at least partly explained by the time use pattern change induced by remittances. To do so, an interaction term per capita remittances received interacted with gender (boy =1)is added in our regression where the dependent variable is time spent on house work, the only time use measure that is significantly affected by remittances [5]. Following our previous approach, only children between ages 5 and 17 are included in the analysis as time use patterns for children beyond 17 years old are not reported. Table 3.11 shows that while the coefficient on the interaction term in the plain fixed effects model is not

[5]The same regressions for time spent on homework and outside work were performed, and results show no significant difference between boys and girls.

statistically significant, it does become significant when endogeneity bias is corrected. The interaction term is positive, which implies that compared to girls in high remittance-receiving families, boys in high remittance-receiving families tend to spend more time on housework. Empirically, this result partly explains why boys' school attendance rate is affected more negatively by remittances compared to girls: as adults migrate out of the country, an under-supply of adult labor occurs. Consequently, boys are brought in to substitute for the adult labor insufficiency on housework, thereby leading to lower school attendance rate.

Table 3.11: Effect of Remittances on Time Use (By Gender)

VARIABLES	(1) Farm Work FE	(2) Farm Work FE, IV
Remittances		
Log Per Capita Remittances	0.0211***	0.0533
	(0.00791)	(0.0978)
Log Per Capita Remittances × Boy	-0.0113	0.211*
	(0.0109)	(0.126)
Personal Characteristics		
Age	0.173***	0.168***
	(0.00790)	(0.00998)
Household Size	0.0342*	0.00673
	(0.0188)	(0.0249)
Age of Household Leader	-0.00116	-0.00175
	(0.00344)	(0.00379)
Gender of Household Leader	-0.0780	0.0350
	(0.106)	(0.126)
Education Level of Household Leader	0.0262	0.0373
	(0.0299)	(0.0324)
House Value	3.33e-08*	2.98e-08
	(1.71e-08)	(1.89e-08)
Ethnicity Control	Yes	Yes
Constant	-0.967***	-1.011***
	(0.273)	(0.302)
Observations	12,413	12,408
Number of ID	4,608	4,608

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

3.5 Conclusion

In a country where labor migration is a prominent phenomenon and where remittances comprise a significant proportion of household income, examining how and to what extent remittance income is allocated for human capital investment is vital. More specifically, ascertaining the impact of remittances on educational expenditure and educational outcomes among children left behind is a development concern that deserves an empirical investigation. The case of the Kyrgyz Republic is a particularly interesting area of study given the critical importance of this issue to the country, as well as its dual experience of development and transition in recent years.

This paper finds that remittances, which constitute a large share of national income, have a negative effect on human capital investment (i.e., educational outcome and expenditure), though the effect is not robust to endogeneity correction. The negative relationship can be at least partially explained by the large positive effect of remittances on expenditure on durable goods. Moreover, this study finds that remittances have a negative effect on educational outcome, that is, on the school attendance rate of children. This can be attributed to the increase of child labor in agricultural/farm work, especially for boys. The main finding of this study calls for actions that mitigate negative impacts as well as to look for ways to incentivize families to invest remittances for education. Intervention programs may be needed to improve financial literacy so that short-term expenditures (such as on durable goods) and long-term expenditures (such as on education and health) can be balanced. Subsidizing farm assets and healthcare may also prove beneficial to effectively utilize remittances for education.

Furthermore, the study shows that the existence of legal and institutional frameworks targeting child labor does not automatically lead to better educational outcomesaddressing gaps in the design and implementation of these frameworks is necessary. The Kyrgyz Government has made advancement in efforts to eliminate the worst forms of child labor, including the ratification of all key international conventions concerning child labor [6] , as well as the establishment of national laws, regulations and policies related to child labor [7]. However, gaps remain in terms of the coverage and comprehensiveness of national frameworks, as well as the inability/inadequacy of laws and regulations to meet international standards. For instance, children are required to attend school only until grade, typically when they reach age 14 or 15. This standard makes children ages 14 and 15 particularly vulnerable to child labor as it is not compulsory for them to attend school but are also not yet legally permitted to work. Noting the high prevalence of child labor in the agriculture sector (i.e., 98.9% of working children aged 10-14 were engaged in agricultural work) [8] , monitoring farm labor hours of school-age children may help ensure appropriate enforcement of child labor laws and encourage school attendance.

Lastly, improving the quality of and access to education and technical/vocational education and training (TVET), could help promote the overall attractiveness of secondary and post-secondary education, and ultimately improve childrens school performance/educational outcomes. In this regard, increasing public expenditure for education, greater investments in quality teaching, and cul-

[6]Including the ILO C. 138, Minimum Age; ILO C. 182, Worst Forms of Child Labor; UN CRC; UN CRC Optional Protocol on Armed Conflict; UN CRC Optional Protocol on the Sale of Children, Child Prostitution, and Child Pornography; and Palermo Protocol on Trafficking in Persons (as cited by the United States Department of Labor, 2017).

[7]Including minimum age for work and hazardous work; prohibition of forced labor, child trafficking, commercial sexual exploitation of children, using children in illicit activities, and military recruitment; compulsory education age; and free public education (as cited by the United States Department of Labor, 2017).

[8]As cited by the United States Department of Labor (2017) based on Understanding Childrens Work Projects analysis of statistics from National Child Labour Survey, 2014.

tivating partnerships with the private sector are integral not only in enhancing the quality and accessibility of the countrys education system but also, and more importantly, address the skills mismatch in the labor market.

References:

[1] Acharya, Chakra P., and Roberto LeonGonzalez. 2014. How do Migration and Remittances Affect Human Capital Investment? The Effects of Relaxing Information and Liquidity Constraints. Journal of Development Studies 50 (3): 444460.

[2] Acosta, Pablo. 2011. School Attendance, Child Labour, and Remittances from International Migration in El Salvador. Journal of Development Studies 47 (6): 913936.

[3] Acosta, Pablo, Pablo Fajnzylber, and Humberto Lopez. 2007. The Impact of Remittances on Poverty and Human Capital: Evidence from Latin American Household Surveys. World Bank Policy Research Working Paper No. 4247.

[4] Adams Jr., Richard H., and Alfredo Cuecuecha. 2010. Remittances, Household Expenditure and Investment in Guatemala. World Development 38 (11): 16261641.

[5] Adams Jr., Richard H. 2006. International Remittances and the Household: Analysis and Review of Global Evidence. Journal of African Economies 15 (2): 396425.

[6] Alcaraz, Carlo, Daniel Chiquiar, and Alejandrina Salcedo. 2012. Remittances, schooling, and child labor in Mexico. Journal of Development Economics 97 (1): 156165.

[7] Amuedo-Dorantes, Catalina, Annie Georges, and Susan Pozo. 2008. Migration, Remittances and Childrens Schooling in Haiti. Institute for the Study of Labor Discussion Paper No. 3657.

[8] Amuedo-Dorantes, Catalina, and Susan Pozo. 2010. Accounting for Remittance and Migration Effects on Childrens Schooling. World Development 38 (12): 17471759.

[9] Amuedo-Dorantes, Catalina, and Susan Pozo. 2006. Migration, Remittances, and Male and Female Employment Patterns. American Economic Review 96 (2): 222226.

[10] Anderson, Kathryn, and Bakhrom Mirkasimov. 2010. Migration, Remittances, and the Human Capital of Children: A Case Study of Tajikistan. Paper presented at the Fourth Southeastern International/Development Workshop, December 3. Atlanta, GA: Federal Reserve Bank and Mercer University.

[11] Ang, Alvin P., Guntur Sugiyarto, and Shikha Jha. 2009. Remittances and Household Behavior in the Philippines. ADB Economics Working Paper Series No. 188.

[12] Antman, Francisca M. 2011. The intergenerational effects of paternal migration on schooling and work: What can we learn from childrens time allocations? Journal of Development Economics 96 (2): 200208.

[13] Bang, James T., Aniruddha Mitra, and Phanindra V. Wunnava. 2016. Do remittances improve income inequality? An instrumental variable quantile analysis of the Kenyan case. Economic Modelling 58: 394402.

[14] Bansak, Cynthia, and Brian Chezum. 2009. How Do Remittances Affect Human Capital Formation of School-Age Boys and Girls? American Economic Review 99 (2): 14548.

[15] Bohme, Marcus H. 2015. Does migration raise agricultural investment?

An empirical analysis for rural Mexico. Agricultural Economics, 46 (2): 211-225.

[16] Bouoiyour, Jamal, Amal Miftah, and El Mouhoub Mouhoud. 2016. Education, male gender preference and migrants' remittances: Interactions in rural Morocco. Economic Modelling 57 (C): 324331.

[17] Bredl, Sebastian. 2011. Migration, remittances and educational outcomes: The case of Haiti. International Journal of Educational Development 31 (2): 162168.

[18] Brown, Richard, Saodat Olimova, and Muhammadi Boboev. 2008. Remittances and Poverty in Central Asia and South Caucasus (Tajikistan): Country Report on Remittances of International Migrants in Tajikistan. Asian Development Bank. Manila: Philippines.

[19] Bucheli, Jos R., Alok K. Bohara, and MatasFontenla. 2018. Mixed effects of remittances on child education. IZA Journal of Development and Migration 8 (1).

[20] Bui, Thi Thanh Nga, Thi Thanh Ngan Le, and Kevin James Daly. 2015. Microlevel impacts of remittances on household behavior: Viet Nam case study. Emerging Markets Review 25: 176190.

[21] Calero, Carla, Arjun S. Bedi, and Robert Sparrow. 2009. Remittances, Liquidity Constraints and Human Capital Investments in Ecuador. World Development 37 (6): 11431154.

[22] Catrinescu, Natalia, Jack Willis, Ian MacAuslan, Valentina Barca, Veronica Sandu, and Martin Rimmer. 2011. Impact of Labour Migration on Children left behind in Tajikistan. United Nations Childrens Fund.

[23] Cebotari, Victor, Melissa Siegel, and Valentina Mazzucato. 2016. Migration and the education of children who stay behind in Moldova and Georgia. International Journal of Educational Development 51 (C): 96107.

[24] Chami, Ralph, Connel Fullenkamp, and Samir Jahjah. 2003. Are Immigrant Remittance Flows a Source of Capital for Development? IMF Working Paper No. 03/189.

[25] Clement, Matthieu. 2011. Remittances and Household Expenditure Patterns in Tajikistan: A Propensity Score Matching Analysis. Asian Development Review 28 (2): 5897.

[26] Cortes, Patricia. 2015. The Feminization of International Migration and its Effects on the Children Left Behind: Evidence from the Philippines. World Development 65 (C): 6278.

[27] Cox-Edwards, Alejandra and Manuelita Ureta. 2003. International migration, remittances, and schooling: evidence from El Salvador. Journal of Development Economics 72 (2): 429461.

[28] de Brauw, Alan, and Scott Rozelle. 2008. Migration and household investment in rural China. China Economic Review 19 (2): 320335.

[29] Demurger, Sylvie, and Xiaoqian Wang. 2016 Remittance and expenditure patterns of the left behinds in rural China. China Economic Review 37: 177190.

[30] Ersado, Lire. 2010. The Global Economic Crisis, Migration, and Remittance Flows to Armenia: Implications for Poverty. Paper presented at the International Workshop on the Economic and Social Impact of Migration, Remittances, and Diaspora, Armenia, June 24-25, 2010. Economic Development Research Cen-

ter, World Bank.

[31] Esenaliev, Damir. 2018. Kyrgyz Republic Country Diagnostic Study: Human Resources. Stockholm International Peace Research Institute [Draft].

[32] Frisancho Robles, Vernica, and Ralph Salvador Oropesa. 2011. International Migration and the Education of Children: Evidence from Lima, Peru. Population Research and Policy Review 30 (4): 591618.

[33] Funkhouser, Ellen. 2013. Using longitudinal data to study migration and remittances. in Vargas-Silva, C. (Ed), Handbook of Research Methods in Migration, London: Edward Elgar Publishing.

[34] Gibson, John, David McKenzie, and Steven Stillman. 2013. Accounting for Selectivity and Duration-Dependent Heterogeneity when Estimating the Impact of Emigration on Incomes and Poverty in Sending Areas. Economic Development and Cultural Change 61 (2): 247280. Global Knowledge Partnership on Migration and Development. https://www.knomad.org/

[35] Gupta, Sanjeev, Catherine A. Pattillo, and Smita Wagh. 2009. Effect of Remittances on Poverty and Financial Development in Sub-Saharan Africa. World Development 37 (1): 104115.

[36] Hagedorn, Annelise, Donghui Wang, and Guangqing Chi. 2017. Remittances and Changing Household Spending Strategies: Evidence from the Life in Kyrgyzstan Study, 20112013.

[37] Halpern-Manners, Andrew. 2011. The effect of family member migration on education and work among nonmigrant youth in Mexico. Demography 48 (1): 7399.

[38] Hanson, Gordon H., and Christopher C. Woodruff. 2003. Emigration and Educational Attainment in Mexico. Mimeo. University of California at San Diego.

[39] Kandel, William and Douglas S. Massey. 2002. The culture of Mexican migration: a theoretical and empirical analysis. Social Forces 80 (3): 9811004.

[40] Karymshakov, Kamalbek and Burulcha Sulaimanova. 2017. Migration impact on left-behind womens labour participation and time-use: Evidence from Kyrgyzstan. World Institute for Development Economics Research. United Nations University.

[41] Karymshakov, Kamalbek, Raziiakhan Abdieva, Burulcha Sulaimanova, and Kadyrbek Sultakeev. 2016. Remittances impact on youth labour supply: evidence from Kyrgyzstan. Working Paper 201605. Partnership for Economic Policy.

[42] Karymshakov, Kamalbek, Raziahan Abdieva, and Burula Sulaymanova. 2014. Workers Remittances and Poverty in Kyrgyzstan.

[43] Kireyev, Alexei. 2006. The Macroeconomics of Remittances: The Case of Tajikistan. IMF Working Paper WP/06/2.

[44] Koska, Onur A., Perihan zge Saygin, Selim aatay, and Andrs Artal-Tur. 2013. International migration, remittances, and the human capital formation of Egyptian children. International Review of Economics and Finance 28 (C): 3850.

[45] Kroeger, Antje and Kathryn H. Anderson. 2014. Remittances and the human capital of children: New evidence from Kyrgyzstan during revolution and financial crisis, 20052009. Journal of Comparative Economics, 42 (3): 770785.

[46] Lall, Somik V., Selod Harris, and Shalizi Zmarak. 2006. Rural-Urban Migration in Developing Countries: A Survey of Theoretical Predictions and Empirical Findings. World Bank Policy Research Working Paper, No 3.915.

[47] Lpez Crdova, Jos Ernesto. 2006. "Globalization, Migration and Development: The Role of Mexican Migrant Remittances." INTAL-ITD Working Paper No. 2652. Inter-American Development Bank.

[48] Lu, Yao, and Donald Treiman. 2007. The Effect of Labor Migration and Remittances on Childrens Education among Blacks in South Africa. California Center for Population Research Working Paper No. 001-07.

[49] Mansuri, Ghazala. 2006. Migration, School Attainment, and Child Labor: Evidence from Rural Pakistan. World Bank Policy Research Working Paper No. 3945.

[50] McKenzie, David, and Hillel Rapoport. 2011. Can migration reduce educational attainment? Evidence from Mexico. Journal of Population Economics 24 (4): 1331-1358.

[51] McKenzie, David, and Hillel Rapoport. 2007. Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. Journal of Development Economics 84 (1): 1-24.

[52] McKenzie, David and Dean Yang. 2010. Experimental Approaches in Migration Studies. World Bank Policy Research Working Paper No. 5395.

[53] Muktarbek kyzy, Akylai, Tchoro Seyitov, and Nurbek Jenish. 2015. Remittances and Expenditure Patterns of Households in the Kyrgyz Republic. Working Paper No. 2. The National Bank of the Kyrgyz Republic.

[54] Munshi, Kaivan. 2003. Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. The Quarterly Journal of Economics 118 (2): 549-599.

[55] Passel, Jeffrey. 2006. Changing dynamics and characteristics of immigration to the United States. Pew Hispanic Center Report.

[56] Quisumbing, Agnes R., and Scott McNiven. 2010. Moving forward, looking back: the impact of migration and remittances on assets, consumption and credit constraints in the rural Philippines. Journal of Development Studies, 46 (1): 91-113.

[57] Salas, Vania B. 2014. International Remittances and Human Capital Formation. World Development 59, (C): 224237.

[58] Stephenson, Andrew V. and Amanda Wilsker. 2016. Consumption Effects of Foreign Remittances in Jamaica. International Advances in Economic Research 22 (3): 309-320.

[59] United States Department of Labor. 2017. Findings on the Worst Forms of Child Labor: The Kyrgyz Republic. https://www.dol.gov/agencies/ilab/resources/reports/child-labor/kyrgyz-republic.

[60] Woodruff, Christopher, and Rene Zenteno. 2007. Migration networks and microenterprises in Mexico. Journal of Development Economics 82 (2): 509-528.

[61] World Bank. 2015. Labor Migration and Welfare in the Kyrgyz Republic (2008-2013).

[62] Yang, Dean. 2008. International Migration, Remittances, and Household

Investment: Evidence from Philippine Migrants Exchange Rate Shocks. Economic Journal 118 (528): 591-630.

[63] Yang, Dean, and HwaJung Choi. 2007. Are Remittances Insurance? Evidence from Rainfall Shocks in the Philippines. The World Bank Economic Review 21 (2): 219-248.

[64] Zhang, Hongliang, Jere R. Behrman, C. Simon Fan, Xiangdong Wei, and Junsen Zhang. 2014. Does parental absence reduce cognitive achievements? Evidence from rural China. Journal of Development Economics 111 (C): 181-195.

CHAPTER 4

KINSHIP, SOCIAL PREFERENCES AND VOTING IN RURAL CHINA: A LAB-IN-THE-FIELD EXPERIMENT

Abstract

Economists have come to understand that human choices are not only driven by self-interest but also "social preferences" – a persons concern over resources allocated to other people. Moreover, such preferences may be affected by the environment in which such choices are made, especially social networks and social pressure. We performed a lab-in-the-field experiment in rural China, where we recruited 162 Chinese farmers to vote in 7 variants of allocation games in randomly assigned groups and with real-world social contacts, with and without pressure. We find that social network and social pressure combined have significant yet heterogeneous effects on social preferences. The source of heterogeneity includes the assignment with in-group or out-group members, membership in dominant lineages, individual characteristics as defined by age and gender, and the degree of kinship between individuals within a social group. Our study not only provides empirical evidence for the social preference theories but also urges policy makers to be careful in choosing an appropriate voting method. In addition, constraining the power of dominant lineage and having better educated villagers more involved in village affairs could be welfare improving.

4.1 Introduction

Economic behaviors of human beings have been modeled on the basic assumption that individuals are rational agents driven exclusively by self-interest, maximizing their own utilities defined in monetary or material terms. However, this assumption is often contradicted by casual observations in history [1] as well as in the real world [2]. In recent years, "social preferences" – a persons concern over resources allocated to others have been introduced to complement self-interest [3]. Parallel to the theoretical modeling are lab experiment results that refute the self-interest assumption and reveal that people exhibit social preferences when making economic decisions (Guth, Schmittberger & Schwarze, 1982; Engelmann and Strobel, 2004; Bolton and Ockenfels, 2006; Messer et al., 2010). A key aspect that is missing from the literature, however, is whether or not such social preferences are dependent on social network (whose interests are involved) and social pressure (how such preferences are elicited). In this paper, we perform a lab-in-the-field experiment where we introduce dyad-level social networks (the lineage) and randomly assignment of pressure treatment into the experimental design. We invite villagers in rural China to bring in their real-world social contacts to play an array of allocation games, with and without social pressure. We find that social preferences are affected the combined force of social network and social pressure. The effects, however, are heterogeneous and vary with lineage dominance, personal characteristics, as well as distance between individuals within a lineage.

The primary objective of our study is to address the incompetence of lab experiments in identifying network-dependent social preferences. In most of the lab experiments, "the subjects enter the laboratory as equals, they do not know

[1] Mass demonstrations to overturn dictatorships in China and Eastern Europe

[2] Blood donation, volunteer activities, etc.

[3] Three major types of social preferences are theoretically modeled: altruism, where people positively value material resources allocated to relevant reference agents (Andreoni, 1989; Cox et al., 2001; modeled as efficiency in Engelmann and Strobel, 2004); inequality aversion, where people prefer an equitable distribution of material resources (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), and reciprocity, where people respond to actions that are perceived to be kind in a kind manner, and to actions that are perceived to be hostile in a hostile manner (Rabin, 1993; Charness and Rabin, 2002)

anything about each other and they are allocated to different roles in the experiment at random", while in reality, "the social context, the saliency of particular agents, and the social proximity among individuals, are all likely to influence reference groups and outcomes"(Fehr and Schmidt, 1999). This is supported by a observational studies on the impact of social network characteristics on cooperative behaviors. These studies have a wide geographical and cultural coverage and show evidence from the United States (Alesina et al., 1997), India (Banerjee, Iyer and Somanathan, 2005), the United Kingdom (Bandiera et al., 2005), Kenya (Miguel and Gugerty, 2005), and China (Miquel et al., 2012). The findings generally indicate that more fragmented social network structure reduces cross-group cooperation and leads to lower level of public goods provision. Therefore, even if a lab experiment finds that people care more about fairness than total welfare (Bolton and Ockenfels, 2006), it may not be the case in the real world because in the real world people are embedded in networks instead of floating around like free atoms.

Acknowledging its importance, more recent lab studies attempt to incorporate social network into their experimental designs. Their approaches of introducing networks, however, are either temporary (on-site assignment of groups within the lab) or categorical (group affiliation instead of dyad level interpersonal relationships). One example of the temporary assignment approach is Chen and Li (2009), where participants that prefer paintings by the same artists are assigned to the same social group. This "group identity" is then reinforced through a few rounds of within-group interactions. This approach, however, fails to incorporate personal histories in the real world and is likely to be rather weak. Guala and Filippin (2005), for example, manipulates the complexity of distributive tasks to show that the induced group identity is merely a framing effect that can be easily displaced by alternative decision heuristics. On the other end of the spectrum is the categorical assignment using existing social groups, for example college fraternities in Kollock (1998) and tribes in Bernhard et al. (2006). They found that group affiliation has a significant effect on distributional preferences – people favor in group members at the expense of outgroup members. The problem with this method is that it homogenizes social ties within a group and mutes the more nuanced effects of interpersonal relationships. Namely, if A and B belong to the same fraternity, they must favor everyone else in the fraternity to the same extent and treat anyone outside the fraternity with the same degree of hatred. Therefore, to more realistically characterize the effect of network on preferences, a dyad-level approach that reflects the mutual, specific, and context-bound interactions between individuals is necessary.

Bringing a community with identifiable dyad-level personal relationships into a behavioral lab is mostly unrealistic due to long travel distance and conflicting schedules of community members. We therefore decide to carry the lab to the field and invite participants to bring in their real-world social contacts to a set of experiments without having to travel from their communities. The design of our experiment (explained in detail in Section 4.3) will allow us to examine the nuanced dependence of social preferences on social network. We also conduct a post-experiment survey to collect information on the specifics of the interpersonal relationship between each pair of participants, which allows us to investigate (1) how participants perceive resources allocated to in-group and out-group members; (2) if members of dominant social groups have different social preference patterns; (3) if social preferences vary with personal characteristics such as gender and age; (4) if variations in social preferences can be explained by social proximity.

Another aspect that is missing from the literature is how social pressure affects social preferences, especially when such pressure interacts with social network effects. It is true that some recent economic experiments incorporate peer pressure as a form of informal sanction and find that disapproval of other agents can increase contribution to public goods (Masclet et al, 2003), and that the allocators in third-party allocation games choose less efficient but more equal distributions when recipients are identifiable (Halali et al, 2017). These lab experiments, however, still recruited independent participants and therefore failed to test how social pressure adds on to social network. Our study, on the other hand, investigates the interaction effect of social pressure and social network by randomly assigning participants into a control group, where anonymity is ensured, and a treatment group, where participants are informed that their choices will be revealed to other participants. This assignment is performed in the social network context where participants play with their social contacts and in the random assignment context where participants play with unknown members. We also interact the social pressure with social proximity between participants to test if the pressure effect depends on degree of kinship.

The current paper is to our knowledge the first lab-in-the-field experiment that embeds dyad level, real-world network structure into allocation games. By observing choices made by any pair of participants in our games and obtaining information on their social interaction in real life, we are able to examine the complex interplay of social preferences and social pressure in the context of the very details of social network. Our study not only serves as an empirical test for the long debated social preference theories but also provides insights at the policy level. Our results show that in any society where social interactions are frequent and personal, the design of a welfare-enhancing policy where collective action is involved (voting, donation, private provision of public goods, etc.) requires a thorough understanding of the social preference patterns of participants and how such preferences interact with the specific network structure and social norms. Without such considerations, identifying an appropriate roadmap of collective action (in the case of voting, policy makers need to choose among secret balloting, show of hands, village meeting, etc.) that yields economically efficient outcomes would be difficult.

The rest of this paper is organized as follows. In section 4.2, we explain the reason why we choose to carry out our experiment in rural China, including the history and characteristics of lineage networks in China; In section 4.3, we introduce our experimental design and procedures. In section 4.4, we present our regression models and empirical results. Section 4.5 concludes.

4.2 Social Network and Social Preferences in the Chinese Context

We choose to implement this experiment in rural China for three reasons, which will be elaborated in the following subsections. First of all, lineages saliently define the basic social network structure in rural China and have long been governing the economic and social behaviors of Chinese villagers (Cohen 2005; Liu and Murphy 2006). Such salience facilitates our recruitment as well as the identification of network effects. Secondly, the rich and observable variations in the patrilineal relationships in the Chinese rural society can be exploited to test how social preferences vary with personal characteristics and with the strength of social ties. Thirdly, the interplay of social network, social pressure, and social preferences is particularly relevant to recent policy changes in village governance in China and other developing countries.

4.2.1 Salience of Lineage Network in Rural China

Social networks exist in numerous forms, but the network in rural china that centers on patrilineal lineage ties is especially salient and deeply entrenched. A lineage is a branch of residents that descend from the same patrilineal ancestors and share the same surname. The patrilineal lineage system has over 3000 years of history form the basis of the Chinese social network (Hu, 2007). Since farmers are attached to their lands, usually the households in a lineage cluster in a settlement for generations, and this long-term connection and repeated interactions make lineage structure within a village rather stable (Coate and Ravallion, 1993). Lineages draw so much on their networks for collective activities that every aspect of village lives revolves around lineages (Cohen 2005; Liu and Murphy 2006; Lu and Tao 2017) [4]. Such stable entrenchment of qualify Chinese villages as an ideal setting to carry out this field experiment.

[4] In terms of economic affairs, lineage groups have taken on governmental functions such as property protection and tax collection (Huang 2008). The provision of village public goods such as elementary schools and irrigation systems has also relied heavily on the donation of major lineage groups (Tsai, 2007; Miquel et. al. 2012). In terms of political affairs, strong lineages influence village elections to ensure members entry into village power operations (Thurston 1998). In some villages, strong lineage networks are mobilized to unite villagers to resist implementation of unpopular birth control policies (Peng 2010). Wary that lineage might counterbalance state power, the Mao government attempted to force peasants to break away from the lineage system and by establishing an administrative village system(He 2003). However, as the economic reforms proceed, the lineage system has undergone a remarkable revival in rural China (Xiao 2001; He 2003) and regained its influence on village affairs.

4.2.2 Heterogeneous Effects of Lineage Network

The advantage of exploiting the lineage network structure is rural China that its rich and identifiable variations allows us to ascertain (1) how participants perceive resources allocated to in-group and out-group members; (2) if members of dominant lineages (big surnames) have different social preference patterns; (3) if social preferences vary with personal characteristics such as gender and age; (4) if variations in social preferences can be explained by social proximity (closer relationship within a lineage).

In-group vs Out-group Allocations

Community in Sociology is often defined as a group of insider people or called insiders-we compared to outsiders-they (Landa 1997: 110, 130). This division of in-group and out-group has been particularly evident in China, where lineage groups tend to be inwardly focused and self-interested and that great inward cohesion [is] gained at the expense of equivalent outward antagonism [5] (Baker, 1979: 1212). It would therefore be interesting to examine how individuals perceive resources allocated to in-group versus out-group members, especially if in-group favors would be offered at the price of hostility towards out-group members (Fukuyama, 2001: 8).

[5] It is also documented in Liu and Murphy (2006) that in many villages in mid 1990s, members of small descent groups were pushed out of their residential villages by larger descent groups and relocating to their ancestral villages. The rationale behind such territorializing actions is that lineage members did not want to allow a share to those who were outsiders

Lineage Dominance

Another interesting feature of the lineage system in rural China is the existence of dominant lineages; that is, a large proportion of residents share the same surname. Relying on the vast of their lands and their prestige in village history, the dominant lineages control the economic and political resources and influence village affairs heavily. Smaller lineage groups, on the other hand, are often suppressed and bullied in the form of name-calling, the vandalism of crops and property and the withholding of irrigation water by the members of the larger kinship groups in their villages (Liu and Murphy, 2006). The display of social preferences may change as the degree of lineage dominance increases (Pan, 2011). This is because in a large lineage group, a members selfish choice may result in more lineage members retaliation. Also the cost of defection potentially rises as the size of the lineage increases, since the deviant can be denied access to a greater amount of resources withheld by the large lineage (He et al, 2017). In our experiment, we will test the how social preferences vary with the degree of lineage dominance.

Personal Characteristics

The lineage ties in rural China are highly individualistic phenomenon in the sense that each tie can only be specified with reference to a particular individual. Hence, an individuals personal characteristics, such as gender and age, can affect his or her position within the network and consequently his or her social preferences. Female members, for example, are positioned inferiorly in this patriarchal system and do not have moral authority as male members do (He 2017). Age is another factor because in the lineage system, seniority equals authority [6]. It is therefore expected that behaviors in our games may depend on personal characteristics of the players.

Network Characteristics

The sacredness of the family tree, the genealogical table that records the human relationship between lineage members, is another feature of the lineage network in rural China. A physical copy of the family tree is kept in the lineage temple, and every member is aware of his position relative to everyone else in the lineage. The strength of dyadic ties in the lineage system depends on social proximity, namely, how close the two agents are in the family tree. The closer agents may therefore display stronger social preferences and may tend to do so at greater costs to the out-group member. We also test this hypothesis in our experiment.

4.2.3 Lineage Network and Rural Democracy

Another reason why we chose rural China for implementation is because the interplay of social network and social preferences has profound policy implications for the democratization movement in recent years. Since late 1990s, the political and economic aspects of village affairs became more and more democratized. Village leaders are elected instead of nominated, and public project proposals need to be voted on before implementation[7]. The outcomes of such collective decision making,

[6]Confucianism supports patrimonial power and emphasizes that everyone should respect their superiors, the young should respect the old and the old should love the young (Hu, 2007)

[7]One example is the introduction of the "One Project, One Review" ("Yi Shi Yi Yi") scheme in 2007, where villagers jointly propose public projects that they wish to implement and vote on the proposal. In addition, welfare allocations are also determined in a more decentralized way in villages nowadays. Villagers gather at village meetings to vote on which households should be

however, largely depend on how individuals perceive the trade-offs between selfish gains and the social optimality when making choices. It has been reported that some villages can hardly ever reach an agreement and hence have never had any public projects implemented. Welfare aids seldom reach the poorest households but end up being allocated to the dominant lineage groups.

One plausible cause, as discussed in the pages before, is that heterogeneous lineage groups often have conflicting interests and tend to only care about in-group members welfare. Social pressure is another plausible factor. An interesting observation we have made from our field research is that villages use various voting methods in village meetings, ranging from a secret ballot to a show of hands. The lack of anonymity may alter the pattern of social preferences displayed in voting, for example making lineage members more fixated on in-group gains, thereby exacerbating the inefficiency of resource allocation. Therefore, we believe carrying out this experiment in the villages in China will help offer valuable policy insights. Moreover, our experimental results can be expanded to other developing countries in the process of democratization, whose formal institutions are weak and informal institutions such as lineages, tribes, or castes guide peoples behavior.

Our experimental design allows us to not only identify the effects of network and pressure on social preferences, but also to ascertain the source of the heterogeneity of such effects (in-group vs out-group, lineage dominance, personal and network characteristics). The identification of preference parameters and network/pressure effects is achieved by manipulating the payoff structure of our allocation games, whereas the identification of the source of heterogeneity is made possible by our post-experiment questionnaire, where we obtain information on demographics of all participants and the degree of kinship between any pair.

provided with the rice or flour allocations from the upper level governments.

4.3 Experimental Design

We designed a series of three-person allocation games and use participants' choices in these games to characterize social preference patterns. To test if voters behave differently when they vote in a network setting (i.e. with their social contacts) or under social pressure, we invite participants to play with anonymous group members and with people from his/her social network, either with social pressure or without. Subsection 4.3.1 discusses the payoff structure and assignment rules of the experiment, and Subsection 4.3.2 describes recruitment and experiment procedure in detail.

4.3.1 Payoff Structure and Assignment Rules

Our experiment consists of 7 variants of a three-person allocation game. The baseline payoff structure of the 7 games is presented in Figure 4.1. Before each game, each participant receives a handout that presents payoffs to the 3 people in his group, with payoff to himself highlighted. Participants are informed that a 3-person group is formed to vote and decide which allocation scheme, out of two alternatives, will be implemented. The voting follows a majority rule. That is, if two or three people in the 3-person group vote yes, each of them will get paid the amount indicated in the Yes column. Otherwise each individual will be paid the amount indicated in the No column. No discussion is allowed. The payoff structure and voting rules have been shown to be incentive compatible, that is, as long as voters have strict preferences over outcomes, voting for ones preferred outcome is the unique trembling-hand perfect equilibrium (Bolton and Ockenfels, 2006).

Figure 4.1: Payoff Structure: Round 1

Game 1 - Round 1						
	Variant 1		**Variant 2**		**Variant 3**	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*
Person 1	2	5	3	5	4	5
Person 2	8	5	7	5	6	5
Person 3	5	5	5	5	5	5

Game 2 - Round 1								
	Variant 1		**Variant 2**		**Variant 3**		**Variant 4**	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Person 1	3	5	9	5	7	5	2	5
Person 2	5	5	5	5	5	5	8	5
Person 3	10	5	4	5	6	5	8	5

The 7 game variants in the baseline case, where the 3 members of each group are randomly assigned before each variant is played (hereafter "Round 1"), are designed to help us understand how preferences for efficiency (total payoff to all 3 players) and equity (fairness of allocation) vary with the pocketbook cost of obtaining them. In the 3 variants of Game 1, for example, overall efficiency is held constant (15 RMB in total for both Yes and No) while the degree of inequality varies across variants. Note that Person 3s self-interest is held constant at 5 RMB between options and across variants, a design that allows us to identify the degree of inequality aversion when no self-interest is involved. For Person 2, obtaining equality requires a sacrifice. We can therefore calculate the proportion of participants willing to sacrifice x units of self-interest in exchange for equality (Person 2, x=1, 2, 3). On the contrary, for Person 1, choosing equality is in line with his self-interest. The 4 variants of Game 2 introduce tension between efficiency and equality. The Yes alternative yields higher overall efficiency (18 RMB) as opposed to the No alternative (15 RMB). The trade-off, however, is that the No alternative is more equal. Note that in the first 3 variants of Game 2, Person 2s payoff is held constant at 5 RMB. This allows us to examine the relative importance of efficiency and equality, when no self-interest is involved. Variations in the payoffs to Person 1 and Person 3, on the other hand, help us identify participants' willingness to sacrifice x units of self-interest and 3 units of overall efficiency in exchange for equality (x=1, 2, 3, 4, 5), and willingness to sacrifice x units of self-interest and equality in exchange for 3 units of efficiency gain (x=1, 2, 3). Note that the way in which the 3 RMB efficiency gain is allocated differs across variants. Variant 3 represents a Pareto gain, Variant 4 represents a majority gain, whereas Variant 1 and 2 represent a majority loss.

We introduce social network in Round 2 by manipulating the 3-person group assignment rule. In the recruitment period, for each session, we recruit 3 original participants (OPs) who are required to come to the experiment with a social contact of his own choice a family member, a relative, or a friend (We term the social contacts Relatives for simplicity). In addition, for each session we recruit 3 villagers who come to the experiment alone (we term them Others). The composition of each session is therefore 3 OPs + 3 Relatives + 3 Others. In Round 2, instead of randomly sampling from the 9 participants in each session to form three 3-person groups as we did in Round 1, we assign any OP and the Relative he brings to the same group, plus one Other. Other remains anonymous throughout the experiment to avoid the possibility that any unobserved personal histories between the pair and the Other could affect their votes. In addition, a new Other is reassigned before each game, so that the pair cannot use the decision made by Other in the previous game to get information on his preference pattern and adjust voting strategy accordingly. We name Round 1 "Random Assignment" and Round 2 "Network Assignment" in our discussion of the experimental results.

The same 7 game variants are played in Round 2. The only difference is in the presentation of payoff tables (shown in Figure 4.2), where players can now identify payoffs to himself, to his Relative, and to Other. In the 3 variants of Game 1, both overall efficiency (total payoff to all 3 players) and within-group efficiency (sum payoffs of OP and Relative) are held constant. This allows us to test how many players are willing to sacrifice x units of self-interest in exchange for within-group equality (OP, x=1, 2, 3), with the out-group members payoff unaffected. This proportion can also be interpreted as the percentage of subjects willing to sacrifice x units in order to help a friend (Altruism). Note that since Others own payoff remains constant across 3 variants, we will be able to observe how the third-party chooses for the within-group allocation when self-interest is not involved. The 4 variants of Game 2, on the other hand, allow us to examine the tension between in-group vs overall efficiency and in-group vs overall equality. In Variant 1, the overall efficiency gain from choosing yes all goes to the out-group member, leaving the in-group efficiency lower in the yes alternative. In Variant 2, the overall efficiency gain all goes to the in-group, leaving the out-group member deprived. Variant 3 represents a Pareto gain both in- and out-group members gain by choosing yes, although the yes allocation is less fair. Variant 4 represents an allocation where the out-group member welfare improves while in-group efficiency remains unaffected. This design allows us to examine how preferences for self-interest, overall efficiency and overall inequality change when participants play with in-group and out-group members.

Figure 4.2: Game Variants: Round 2

Game 1 - Round 2

	Variant 1		Variant 2		Variant 3	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*
Relative	2	5	3	5	4	5
OP	8	5	7	5	6	5
Other	5	5	5	5	5	5

Game 2 - Round 2

	Variant 1		Variant 2		Variant 3		Variant 4	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
Relative	3	5	9	5	7	5	2	5
OP	5	5	5	5	5	5	8	5
Other	10	5	4	5	6	5	8	5

Another dimension of our design, in addition to Random Assignment vs Network Assignment, is the social pressure treatment. While each participant plays in both rounds, the same participant will be assigned into either the Control group (no pressure) or the Treatment group (with pressure), not both. The social pressure treatment aims to test if voters reveal different patterns of social preferences when they are aware that their choices are visible to others, mirroring the potential impact of lack of anonymity in village democracy – for example the show of hands instead of secret ballot. In terms of implementation, we inform participants at the beginning of the session that their votes will be revealed to group members after all games are completed. We choose to reveal the votes at the very end of the experiment instead of after each game so that participants cannot judge the type of their group members and adjust strategically in the subsequent games. The combination of assignments and treatments are summarized in the table below.

Figure 4.3: Assignment-Treatment Combinations

	Control (Without Social Pressure)	**Treatment** (With Social Pressure)
Round 1 (Random Assignment)	3 participants randomly drawn to form a group (changes every game). Votes kept secret throughout experiment.	3 participants randomly drawn to form a group (changes every game). Votes revealed within group after all games end.
Round 2 (Network Assignment)	OP + Relative + 1 randomly assigned "Other" (changes every game). Votes kept secret throughout experiment.	OP + Relative + 1 randomly assigned "Other" (changes every game). Votes revealed within group after all games end.

4.3.2 Experiment Procedure

The formal experiment was implemented in July 2017 in 4 randomly selected villages in Linyi, Shandong Province in China. Before the formal experiment, we performed a pilot experiment in January 2017 in Shandong for test purposes only. Data from the pilot experiment is not used in this paper. The experiment design and consent process was reviewed and approved by the Internal Review Board (IRB) at Cornell University. We obtained oral consent from all subjects for their participation in the experiment. All experiment materials, including payoff cards and questionnaires, were presented in Chinese. Experiment instructions were read to participants in Linyi dialect to ensure understanding.

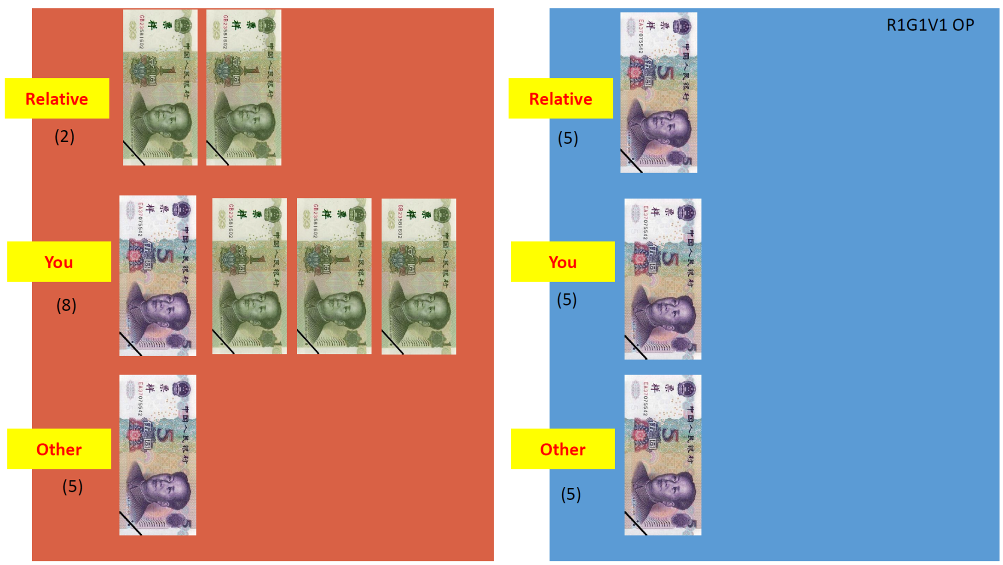
We held 18 experiment sessions with 9 subjects in each session (162 participants in total). For recruiting, we first contacted the leader of each village to explain the purpose of our study and to obtain the leaders consent. The village leader then sent out our recruitment materials (age requirement, time, location, expected monetary payoff ect., see Appendix C) to the entire village to invite voluntary participation. Summary statistics of participants' personal characteristics are presented in the table below. Approximately 2/3 of participants were female, plausibly due to the fact males were working in the field during most of our sessions. Age of participants ranges from 20 to 56, with an average of 41 years old. The participants on average completed 8 years of education, which is equivalent to a sophomore in junior high school. “Name Percentage” is a variable we use to capture the dominance of the lineage group each participant belongs to. For example if a participant has surname “Wang” and 80% of the village population shares the same surname, then his “Name Percentage” is indicated as 0.8. We obtained this information from village leaders and confirmed with randomly selected participants in each village. A by-village breakdown of the summary statistics is provided in Appendix C.

Table 4.1: Summary Statistics

	Mean	Standard Deviation
Female Percentage	66.7%	
Age	41.36	10.38
Years of Education	7.929	2.550
Monthly Income (RMB)	1922.1	2882.5
Name Percentage	0.213	0.280
Observations	162	

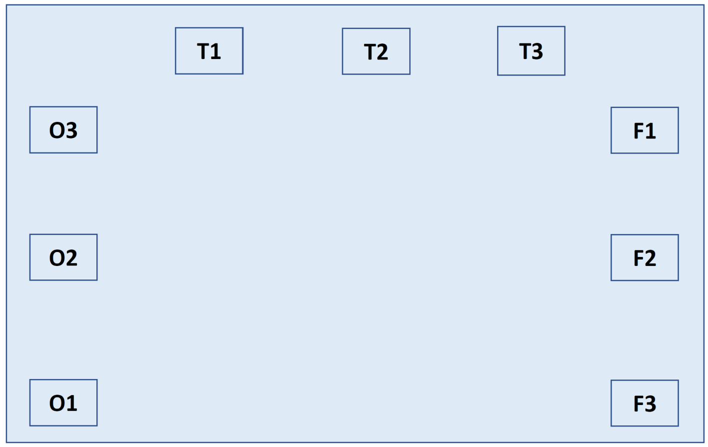
In the pilot experiment, we presented the 7 game variants in table format, as shown in Figure 4.1 and Figure 4.2. However, due to the low education level and limited literacy, a majoriity of our participant had trouble reading and comprehending tables and numbers. We hence modified the presentation in our formal experiment. We use a red and a blue “plate” to represent the yes and no alternatives, as shown in Figure 4.4, with images of the Chinese currency (RMB) on each plate to visualize the payoffs. The yellow labels on the edges of plates indicates ownership of allocation, namely, to whom the money will be allocated (for example, You, Member 2, Member 3 in Round 1, You, Your Relative, Other in Round 2). The small numbers in black under the labels are numerically equivalent to the amount of RMB presented in each row. We included them to help participants double check their counting. Participants in the formal experiment voted red or blue instead of yes or no.

Figure 4.4: Colored Plates to Present Payoffs



9 subjects enter the experiment room and sit in designated seats (Figure 4.5, O=OP, F=Relative, T=Other). The distance between seats was set to be at least 1.5 meters so that subjects cannot communicate or peek each other's vote.

Figure 4.5: Colored Plates to Present Payoffs



The same script of experimental instructions (in Appendix) was read out to participants in Linyi dialect by the same experimenter at every session to ensure consistency. The experimental instruction is essentially the same as described in the previous subsection. At the beginning of Round 1, the main experimenter informs subjects that 3 members will be selected at random to form a group (3 groups in total). Group members are shuffled and reassigned at the beginning of each new game and the identity of group members will not be revealed. She then proceeds to explain the distribution schemes and the majority rule. In the Control sessions (Session 1, 2, 5, 6, 9, 10, 13, 14, 17, 18), participants are informed that their votes will never be made public whereas in the Treatment sessions (Session 3, 4, 7, 8, 11, 12, 15, 16), participants are warned ahead of time that after all games are finished we will reveal everyones votes to group members. The experimenters then distribute payoff plates and instruct participants how to read their own payoffs and other participants payoffs. The majority rule is explained again and confirmation questions are asked to ensure correct understanding. For example, "if you vote red and the other two members vote blue, what would Person 3's payoff be for this game". For each game, we give participants 1 minute to consider before we go collect votes. Participants are instructed to point at the color of their choice instead of stating aloud so that votes remain unknown to other participants. Experimenter writes down votes on a data sheet and uses majority rule to decide which allocation to implement. Corresponding payments are made immediately to participants to provide monetary incentives. Round 2 proceeds the same way as Round 1, with the only exception that participants are informed that the pair (OP + Relative) will always be in the same group while the third member will be randomly assigned before each game and will remain unidentifiable.

After all games are completed, participants are required to fill out a questionnaire about their demographic information. The pairs (OPs and Relatives) are asked to complete an additional set of questions on their relationships, frequency of interaction, etc. Each session, including the survey completion, takes between 45 minutes and 1 hour. The total payoff to each participant ranges from 50 to 90 RMB with mean payoff targeted at 75 RMB, which is equivalent to an average participant's daily income. We believe this provides sufficient financial incentive to participants and ensures that they take their decisions seriously.

4.4 Results

4.4.1 Descriptive Results

In this sub-section, we provide a game-by-game description of participants' social preferences patterns exhibited in the experiment. Due to limited space, we include tables (t-test results) in Appendix C (Part B). Entries of the tables are percentage of players voting No. The more rigorous econometric analysis results will be presented in sub-section 4.4.2.

In all the 3 variants of Game 1 where efficiency is held constant at 15RMB between Yes and No, the majority of participants (over 50% in all variants) chose the fair allocation regardless of treatment. In addition, more than half of participants were willing to sacrifice self-interest (1 RMB, 2 RMB and 3 RMB) to achieve a fair allocation. In terms of treatment effect, in Round 1 (Random Assignment) social pressure does not alter behaviors in a statistically significant manner. In Round 2 (Network Assignment), however, subjects are more willing to make sacrifices when they are under social pressure. In other words, subjects display more pro-social preferences when they are aware that their choices will be revealed to people in their social networks. The motive of such behavior changes – if it is due to stronger preferences for overall equality or within-group equality – will be tested in regression analyses.

The 4 variants of Game 2 allocate the 3 RMB overall efficiency gain differently. In Game 2 Variant 1, the efficiency gain goes exclusively to the out-group member while the in-group members lose. In the Round 1 (Random Assignment)- Control (No Pressure) situation, the proportion of Person 2 players (no self-interest involved) choosing efficiency and the proportion choosing fairness are almost equal. This proportion changed significantly when Person 2 players are informed that they will be paired with their partners (Round 2). Without social pressure, 27% more of them chose the option that increases in-group total payoff, although it results in a larger loss in overall efficiency.

In Game 2 Variant 2, the efficiency gain is allocated to in-group members exclusively. With self-interest neutral across options, the majority of Person 2 players chose efficiency over fairness most of the time. This ratio is only reversed in the Round 1 (Random Assignment)- Treatment (Pressure) situation, where 25% more of Person 2 Players chose fairness over efficiency. Interestingly, 20% more of the out-group members voted No when they are informed the other two are pairs, a sign that the out-group members dislike allocations where the paired players gain more.

The Yes option of Game 2 Variant 3 represents a Pareto gain. Most of our subjects realize this and played accordingly. The large majority of players (67%-87%) chose efficiency over equality, and this strong preference for Pareto efficiency gain is consistent across treatments and across roles. The only exception is that Person 1 players voted more for equality when they play with their partners under pressure. However even after this change, the proportion of Person 1 players choosing efficiency over equality is still as high as 2/3.

Finally, Game 2 Variant 4 represents an efficiency gain that goes to the out-group member while keeping in-group efficiency constant. In this case, in-group members (Person 1 and Person 2) consistently preferred the fair allocation, unwilling to let the efficiency gain go to the out-group member. When Person 1 players are informed that Person 2 is their Relatives, they make significantly more sacrifices to let Person 2 earn more.

In summary, when efficiency is held constant, subjects displayed strong distaste towards inequality and are willing to sacrifice self-interest for equality even in the Random Assignment. Social pressure, when efficiency is constant, only works under Network Assignment. When there is an efficiency gain, however, behavioral change depends heavily on how the efficiency gain is allocated in-group vs. out-group. We also observe that when subjects play under Network Assignment, they become fixated on within-group gains even when doing so leads to an overall efficiency loss, and such tendency becomes more prominent when social pressure is present. A Pareto gain, on the other hand, is strongly preferred regardless of assignment and treatment.

4.4.2 Regression Analysis

Social Preferences, Network Effects, and Pressure Effects

The rich variation in payoff structure of the 7 game variants allows us to identify parameters that characterize preference for self-interest, efficiency, and inequality aversion, since these 3 components have been found to foster sharply different conduct (Rabin 1993, Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Charness and Rabin 2002, Engelmann and Strobel 2004, Fehr and Schmidt 2006). While self-interest and efficiency have been consistently defined in the existing literature as payoff to oneself and the sum of payoffs to all members respectively, the definition of equality has been under debate. We include three measures of inequality aversion – FS, ERC, and MaxiMin – in our regression analyses since they not only form the basis of theoretical work but also are the most widely employed ones in empirical tests (Engelmann and Strobel, 2004; Messer et al, 2010).

Attempting to explain cooperative behavior by a single simple model, Fehr and Schmidt (1999, henceforth FS) model fairness as self-centered inequality aversion; namely, people do not care about inequity per se that exists among others but are only interested in the fairness of their own material payoff relative to the payoff of others. Mathematically, the utility that person i generates from the game outcome is written as:

$$U_i(\pi) = \pi_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} max\{\pi_j - \pi_i, 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i} max\{\pi_i - \pi_j, 0\}$$

$$\alpha_i \geq \beta_i, \;\; 0 \leq \beta_i < 1$$

where π_i is the payoff to i himself, π_j is the payoff to any other player j, and n is the total number of players. In this framework the two max terms are mutually exclusive: α_i measures the disutility from other members being better off (Namely, when i is Behind), while β_i measures the disutility from other members being worse off (namely, when i is Ahead), $\alpha_i \geq \beta_i$ therefore captures the idea that a player suffers more from inequality that is to his disadvantage. They then tested the model on a wide array of games such as ultimatum games, market games, and public goods games. The result is mixed in that in some games (ultimatum, public goods with punishment) more cooperative behaviors are observed whereas in other games (market, public goods without punishment) subjects behave more selfishly. They therefore conclude that the distribution of social preferences depends heavily on the strategic environment of the games.

Bolton and Ockenfels (2000, Equity, Reciprocity, and Competition, henceforth ERC) organized a large set of laboratory games with a simple model constructed on the premise that people are motivated by both their pecuniary payoff and their relative payoff standing. In particular, individual is utility function is specified as:

$$U_i = U_i(\pi_i, \sigma_i)$$

$$c = \sum_j \pi_j$$

$$\sigma_i = \sigma_i(\pi_i, c, n) = \begin{cases} \pi_i / c & \text{if } c > 0 \\ 1/n & \text{if } c = 0 \end{cases}$$

$$U_i(c\sigma_i, \sigma_i) = \alpha_i c\sigma_i - \frac{\beta_i}{n}(\sigma_i - \frac{1}{n})^2$$

$$\alpha_i \geq 0, \beta_i > 0$$

Namely, the utility is a function of payoff to i himself (π_i) and the share he gets in the total payoff (σ_i). Utility is assumed to be strictly concave in the share argument, with a maximum around the allocation at which ones own share is equal to the average share. This implies that the egalitarian division is most preferred. In essence, the further the allocation moves from player i receiving an equal share, the higher the loss from the comparative effect. A players type is characterized by α_i/β_i, the ratio of weights that are attributed to the pecuniary and relative components. The authors then fit data from various games (dictator, gift-exchange, Prisoners Dilemma, etc) and find that ERC equilibrium predicts peoples strategic behavior quite well.

Charness and Rabin (2002, henceforth MaxiMin) measures inequality by the payoff to the worst-off person. They designed an array of new experimental games to determine whether subjects are more concerned with increasing social welfare sacrificing to increase the payoffs for all recipients, especially low-payoff recipients than with reducing differences in payoffs. The multi-person MaxiMin model is specified as:

$$U_i(\pi_1, \pi_2, ..., \pi_N) = (1-\lambda)\pi_i + \lambda[\delta \cdot min(\pi_1, \pi_2, ..., \pi_N) + (1-\delta)(\pi_1 + \pi_2 + ... + \pi_N)]$$

$$\lambda \in [0, 1]$$

That is, subjects like money but also prefer Pareto-improvements. By positing a concern for efficiency, this model helps explain why many subjects make inequality-increasing sacrifices because these choices are Pareto-improving and inexpensive.

A number of empirical studies in the 2000s employ new game designs to compare performances of the above social preference models. Engelmann and Strobel (2004), for example, presents a set of three-person one-shot distribution experiments to examine the importance of efficiency and inequality aversion in decision-making. They also compare the relative performance of the three models mentioned above. They find that the multi-person MaxiMin model can rationalize most of the data while neither ERC or FS can explain important patterns. In response, Bolton and Ockenfels (2006) challenged the Engelmann-Strobel design since the decision makers self-interest remained unaffected in most cases namely no sacrifice was necessary. They then performed additional experiments to show that *willingness to pay* for efficiency is substantially lower than it is for equity. They also look more closely at the role of procedural equity by manipulating the role assignment rule in a majority voting game, and find that equal opportunity procedures can soften the tension between equality and efficiency. Messer et al (2010) uses a Random Price Voting Mechanism (RPVM) to elicit social preferences in a referenda experiment to find that the two equity based models ERC and FS are not supported by the data while MaxiMin performs relatively well. They conclude that a social efficiency motive may lead to economically meaningful deviations from selfish voting choices and increase the likelihood that welfare-enhancing programs are implemented.

We use participants' choices in each game and the associated payoff structures to estimate preference parameters in the above 3 models respectively. Specifically, we fit the ERC, FS, Maximin utility functions using a binary logit regression as in Charness and Rabin (2002) and Engelmann and Strobel (2004). Assuming that all individuals choose yes or no to maximize their utilities, the probability of individual i voting yes can be written as:

$$P(yes_i) = \frac{e^{U_i(yes_i)}}{e^{U_i(yes_i)} + e^{U_i(no_i)}}$$

Utility is assumed to be separately additive in three elements: self-interest, social efficiency, and inequality (using ERC, FS and MaxiMin one at a time). Note that when empirically applying the theoretical model specified in Fehr and Schmidt (1999), multi-colinearity is unavoidable since $FS\ Ahead = FS\ Behind$ $+\ Efficiency$ - $3\times$ $Self - Interest$. To overcome this problem, we follow Engelmann and Strobel (2004) by using two approaches. In a first approach, we exclude self-interest (we term it the FS model when presenting regression results). In a second approach, we employ Engelmann and Strobel's strict version by specifying $FSstrict\ =\ FS\ Ahead\ +\ FS\ Behind$. This new measure is essentially an aggregation of inequality where equal weights are assigned to disadvantageous and advantageous inequality (we term it the FS Strict model when presenting regression results).

Table 4.2: Utility Specification in Regressions

	Utility Function of Person i
ERC	$\alpha_i^{SELF,ERC}\pi_i + \beta_i^{EFF,ERC}\sum_j \pi_j + \beta_i^{ERC}\lvert\pi_i - \frac{1}{N}\sum_{j=1}^{N}\pi_j\rvert$
FS	$\beta_i^{EFF,FS}\sum_j \pi_j + \beta_i^{FS,AHEAD}\frac{1}{N-1}\sum_{j\neq i} Max(\pi_i - \pi_j, 0) + \beta_i^{FS,BEHIND}\frac{1}{N-1}\sum_{j\neq i} Max(\pi_j - \pi_i, 0)$
FS Strict	$\alpha_i^{SELF,FSst}\pi_i + \beta_i^{EFF,FSst}\sum_j \pi_j + \beta_i^{FSst}\frac{1}{N-1}[\sum_{j\neq i} Max(\pi_j - \pi_i, 0) + \sum_{j\neq i} Max(\pi_i - \pi_j, 0)]$
Maximin	$\alpha_i^{SELF,MM}\pi_i + \beta_i^{EFF,MM}\sum_j \pi_j + \beta_{i1}^{MM}\cdot Min(\pi_1,..., \pi_N)$

The design of our experiment, especially the Assignment-Treatment Combinations as shown in Figure 4.3, allows us to identify two effects. Network effect characterizes if participants display different degree of preferences for self-interest, efficiency, and equality when they are assigned to the same group as their social contacts, with or without pressure. Pressure effect, on the other hand, characterizes if participants display different degree of preferences for self-interest, efficiency, and equality when they are under social pressure, with or without social contacts in their groups. Mathematically, the utility function used to identify network effect, without social pressure ($P = 0$) and with social pressure ($P = 1$) is specified as:

$$U_i(P=0) = \alpha^{N0}_{i,P=0}\pi_i + \beta^{N0}_{i,P=0}\sum_j \pi_j + \gamma^{N0}_{i,P=0}d_i$$

$$+ \alpha^{N1}_{i,P=0}\pi_i \cdot \mathbb{1}_N(i) + \beta^{N1}_{i,P=0}\sum_j \pi_j \cdot \mathbb{1}_N(i) + \gamma^{N1}_{i,P=0}d_i \cdot \mathbb{1}_N(i)$$

$$U_i(P=1) = \alpha^{N0}_{i,P=1}\pi_i + \beta^{N0}_{i,P=1}\sum_j \pi_j + \gamma^{N0}_{i,P=1}d_i$$

$$+ \alpha^{N1}_{i,P=1}\pi_i \cdot \mathbb{1}_N(i) + \beta^{N1}_{i,P=1}\sum_j \pi_j \cdot \mathbb{1}_N(i) + \gamma^{N1}_{i,P=1}d_i \cdot \mathbb{1}_N(i)$$

where π_i represents self-interest, $\sum_j \pi_j$ represents the efficiency measure, d_i represents any inequality measure (ERC, FS, FS Strict, or MaxiMin), and $\mathbb{1}_N(i)$ is an indicator function that takes value 1 if the choice made by individual i is observed from the Network Assignment. We are interested in testing which of the two network effects – network effect without social pressure (which we term Network Effect A, identified as $\alpha^{N1}_{i,P=0}$, $\beta^{N1}_{i,P=0}$, and $\gamma^{N1}_{i,P=0}$) and network effect with social pressure (which we term Network Effect B, identified as $\alpha^{N1}_{i,P=1}$, $\beta^{N1}_{i,P=1}$, and $\gamma^{N1}_{i,P=1}$) – is statistically significant.

Similarly, the utility function used to estimate pressure effect, in the Random Assignment ($N = 0$) and Network Assignment ($N = 1$), is specified as:

$$U_i(N=0) = \alpha^{P0}_{i,N=0}\pi_i + \beta^{P0}_{i,N=0}\sum_j \pi_j + \gamma^{P0}_{i,N=0}d_i$$

$$+ \alpha^{P1}_{i,N=0}\pi_i \cdot \mathbb{1}_P(i) + \beta^{P1}_{i,N=0}\sum_j \pi_j \cdot \mathbb{1}_P(i) + \gamma^{P1}_{i,N=0}d_i \cdot \mathbb{1}_P(i)$$

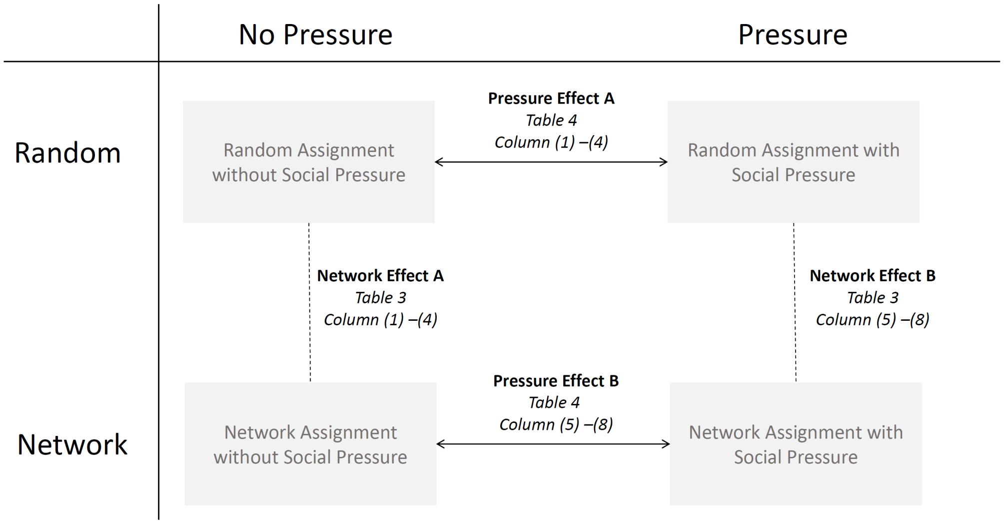
$$U_i(N=1) = \alpha^{P0}_{i,N=1}\pi_i + \beta^{P0}_{i,N=1}\sum_j \pi_j + \gamma^{P0}_{i,N=1}d_i$$

$$+ \alpha^{P1}_{i,N=1}\pi_i \cdot \mathbb{1}_P(i) + \beta^{P1}_{i,N=1}\sum_j \pi_j \cdot \mathbb{1}_P(i) + \gamma^{P1}_{i,N=1}d_i \cdot \mathbb{1}_P(i)$$

where $\mathbb{1}_P(i)$ is an indicator function that takes value 1 if the choice made by individual i is observed from the Treatment Group (with social pressure). We are interested in testing which of the two pressure effects – pressure effect under Random Assignment (which we term Pressure Effect A, identified as $\alpha^{P1}_{i,N=0}$, $\beta^{P1}_{i,N=0}$, and $\gamma^{P1}_{i,N=0}$) and pressure effect under Network Assignment (which we term Pressure Effect B, identified as $\alpha^{P1}_{i,N=1}$, $\beta^{P1}_{i,N=1}$, and $\gamma^{P1}_{i,N=1}$) – is statistically significant.

The above 4 effects are summarized in Table 4.3 and Table 4.4. Each 4 columns present a set of results using the 4 inequality measures respectively – ERC, FS Strict, FS, and Maximin. Column (1) - (4) of Table 4.3 examines Network Effect without social pressure (Network Effect A); Column (5) - (8) of Table 4.3 examines Network Effect under social pressure (Network Effect B); Column (1) - (4) of Table 4.4 examines Pressure Effect in random assignment (Pressure Effect A); Column (5) - (8) of Table 4.4 examines Pressure Effect in network assignment (Pressure Effect B). The logic of our result presentation is summarized in Figure 4.6. We also performed a balance covariates check to make sure the demographics of participants in the control and treatment groups do not differ significantly. The result is presented in the Appendix. Age is the only covariate that differs across control and treatment, and the difference is rather weak (at p=0.1). We are therefore confident to say that the treatment assignment is indeed random.

Figure 4.6: Regression Tables Illustration



The upper panel of the first 4 columns of Table 4.3 characterizes the baseline (no social pressure, no network effects) social preference patterns displayed by our participants. Results are robust across the 4 model specifications in the sense that all social preference parameters are statistically significant and are of the expected signs. Subjects strongly prefer ($p<0.01$) higher payoffs to themselves (self-interest) and higher total payoffs to the group (efficiency). In the mean time, players display significant inequality aversion regardless of the measurement used. We also observe that the FS Behind parameter is (in absolute value) twice the size of the FS Ahead parameter, meaning that subjects suffer more from inequality that is to their disadvantage. This is in line with the theoretical prediction in Fehr and Schmidt (1999). The lower panel of the first 4 columns reveals that when participants are not under social pressure, network effect is statistically insignificant. In other words, when participants are informed that their votes will not be made public, playing with their relatives does not change their preference patterns.

The upper panel of columns (5)-(8) of Table 4.3 characterizes social preference patterns when participants play in Random Assignment under social pressure. Very similar pattern is revealed here as in the baseline case. The lower panel of columns (5)-(8), however, shows very strong Network Effect. That is, when participants are aware that their votes will be revealed to other players, assigning them to play with their relatives will significantly alter their social preference patterns. The change is particularly obvious in inequality aversion, in that players display much stronger distaste for unfair distributions and stronger willingness to help the worst-off person.

Table 4.3: Network Effects: Without and With Social Pressure

	(1) No Pressure ERC	(2) No Pressure FS Strict	(3) No Pressure FS	(4) No Pressure MM	(5) Pressure ERC	(6) Pressure FS Strict	(7) Pressure FS	(8) Pressure MM
Self-Interest	0.163***	0.156***		0.158***	0.231***	0.226***		0.229***
	(0.0407)	(0.0406)		(0.0410)	(0.0471)	(0.0470)		(0.0476)
Efficiency	0.206***	0.258***	0.305***	0.189***	0.170***	0.233***	0.300***	0.157***
	(0.0538)	(0.0569)	(0.0554)	(0.0471)	(0.0602)	(0.0639)	(0.0624)	(0.0529)
ERC	-0.333***				-0.331***			
	(0.0593)				(0.0678)			
FS		-0.122***				-0.127***		
		(0.0194)				(0.0222)		
FS Ahead			-0.137***				-0.0978*	
			(0.0447)				(0.0499)	
FS Behind			-0.351***				-0.404***	
			(0.0494)				(0.0579)	
MM				0.399***				0.403***
				(0.0543)				(0.0615)
Self × Network	-0.0556	-0.0565		-0.0586	-0.0254	-0.0396		-0.0603
	(0.0576)	(0.0575)		(0.0580)	(0.0704)	(0.0693)		(0.0678)
Efficiency × Network	0.0147	0.0317	0.0109	0.0165	0.135	0.157*	0.139	0.0391
	(0.0763)	(0.0810)	(0.0787)	(0.0669)	(0.0878)	(0.0937)	(0.0916)	(0.0758)
ERC × Network	-0.0449				-0.283***			
	(0.0841)				(0.103)			
FS × Network		-0.0211				-0.0915***		
		(0.0277)				(0.0336)		
FS Ahead × Network			-0.0789				-0.206***	
			(0.0640)				(0.0738)	
FS Behind × Network			-0.00177				-0.155*	
			(0.0700)				(0.0877)	
MM × Network				0.0637				0.150*
				(0.0778)				(0.0901)
Observations	2,520	2,520	2,520	2,520	2,016	2,016	2,016	2,016

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

The upper panel of the first 4 columns of Table 4.4 is exactly the same as that of Table 4.3 – it summarizes the baseline (no social pressure, no network effects) social preference patterns. In addition, all parameters in the lower panel of the first 4 columns are statistically insignificant, indicating that when group members are randomly assigned and anonymous, adding social pressure does not alter participants' social preference patterns.

The upper panel of columns (5)-(8) of Table 4.4 presents social preference patterns when participants play in Network Assignment without social pressure. Consistent with results shown in all the upper panels in the previous discussion, all social preference parameters are significant and are of the expected signs. The lower panel of columns (5)-(8) shows statistically significant Pressure Effect in most of our specifications. Participants display stronger inequality aversion (in ERC, FS Strict, and FS) when they are aware that their votes will be revealed to their relatives.

Table 4.4: Social Pressure Effects: Random vs Network Assignment

	(1) Random ERC	(2) Random FS Strict	(3) Random FS	(4) Random MM	(5) Network ERC	(6) Network FS Strict	(7) Network FS	(8) Network MM
Self-Interest	0.163***	0.156***		0.158***	0.107***	0.0996**		0.0998**
	(0.0407)	(0.0406)		(0.0410)	(0.0407)	(0.0408)		(0.0409)
Efficiency	0.206***	0.258***	0.305***	0.189***	0.221***	0.290***	0.316***	0.205***
	(0.0538)	(0.0569)	(0.0554)	(0.0471)	(0.0541)	(0.0576)	(0.0558)	(0.0475)
ERC	-0.333***				-0.378***			
	(0.0593)				(0.0597)			
FS		-0.122***				-0.143***		
		(0.0194)				(0.0198)		
FS Ahead			-0.137***				-0.216***	
			(0.0447)				(0.0458)	
FS Behind			-0.351***				-0.353***	
			(0.0494)				(0.0495)	
MM				0.399***				0.463***
				(0.0543)				(0.0557)
Self × Pressure	0.0684	0.0702		0.0710	0.0986	0.0871		0.0693
	(0.0623)	(0.0621)		(0.0628)	(0.0663)	(0.0652)		(0.0633)
Efficiency × Pressure	-0.0359	-0.0248	-0.00503	-0.0319	0.0845	0.101	0.123	-0.00935
	(0.0807)	(0.0856)	(0.0835)	(0.0708)	(0.0838)	(0.0896)	(0.0872)	(0.0722)
ERC × Pressure	0.00158				-0.237**			
	(0.0901)				(0.0983)			
FS × Pressure		-0.00468				-0.0751**		
		(0.0295)				(0.0321)		
FS Ahead × Pressure			0.0391				-0.0884	
			(0.0670)				(0.0711)	
FS Behind × Pressure			-0.0531				-0.206**	
			(0.0762)				(0.0824)	
MM × Pressure				0.00402				0.0899
				(0.0820)				(0.0862)
Observations	2,268	2,268	2,268	2,268	2,268	2,268	2,268	2,268

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

In summary, Table 4.3 and Table 4.4 combined reveal that in addition to self-interest, participants consistently display preference for social efficiency and distaste for inequality. Such social preferences are intrinsic and robust against alternative specifications. Moreover, social network or social pressure alone cannot alter behaviors. Social preference patterns, in particular inequality aversion, change only when the two forces are combined.

One thing to note is that the above analyses have pooled choices made by all participants, including the pair and the Other. One may suspect that by doing so, preference patterns of the Others may have weakened Network Effects and Pressure

Effects, since Others came alone and are hence less susceptible to judgment made by their social contacts. As a robustness check, we excluded Others from our analyses and only kept choices made by the pairs. The results and conclusions are unchanged.

From the baseline parameters we are able to calculate willingness-to-pay (WTP) for "social goods", namely how much money (in RMB) a person is willing to sacrifice in order to achieve efficiency and equality. The WTP measures are calculated as the ratio of the social preference parameter and self-interest parameter, namely,

$$WTP_{EFF} = MRS_{Self,EFF} = \frac{MU_{EFF}}{MU_{Self}} = \frac{\beta_{EFF}}{\beta_{Self}}$$

$$WTP_{EQ} = MRS_{Self,EQ} = \frac{-MU_{EQ}}{MU_{Self}} = \frac{-\beta_{EQ}}{\beta_{Self}}$$

The results are reported in Figure 4.7. We observe that except for FS, all willingness-to-pay measures are greater than one, meaning that on average participants are willing to give up more than one RMB in exchange for one-RMB increase in efficiency, or one-unit reduction (in squared terms) in equality, or one-RMB increase in the payoff to the worst of person. In other words, social preferences are not only present but also large in magnitude.

Figure 4.7: Willingness to Pay for Social Goods

	ERC	**FS**	**MM**
Efficiency	1.264	1.654	1.196
ERC	2.043		
FS		0.782	
MM			2.525

In-group vs Out-group Allocations

As introduced in Section 4.2, one key feature of the Chinese lineage networks is the clear division of "insiders-we" and "outsiders-they". In this subsection, we examine how people in a network setting perceive resources allocated to in-group members (relatives) as opposed to out-group members (Others).

We introduce three new variables: *In-group Efficiency*, which measures the total payoff to the pair ($\pi_{OP} + \pi_{Relative}$); *In-group Inequality*, which measures the absolute difference in payoffs between the pair ($|\pi_{OP} - \pi_{Relative}|$); and *Distance*, which measures how far the Other's payoff is away from the average payoff of the pair($|\pi_{Other} - \frac{1}{2}(\pi_{OP} + \pi_{Relative})|$). We only use data from the Network Assignment, because the division of in-group and out-group does not exist in Random Assignment. In addition, we separate observations of the pairs (in-group members) and the Others (out-group members), because they may behave differently facing the division and social pressure. Specifically, utility function of the in-group members is specified as a separately additive function of 4 elements: self-interest, overall efficiency, in-group efficiency, and in-group inequality; utility function of the out-group members is specified as a separately additive function of 3 elements: self-interest, overall efficiency, and distance. The results are presented in Table 4.5.

The upper panel of Table 4.5 presents preferences for in-group measures in the network setting without social pressure. The pairs (in-group members) do not seem to be concerned with self-interest nor overall efficiency in this case. Since in-group efficiency is included in the regression, this can also be interpreted as their disregard of the payoffs to the out-group member. Instead, all that they care about are in-group measures – in-group efficiency and in-group inequality. Such fixation on in-group measures is exacerbated by social pressure, as we observe from the lower panel of column (1). Imposing social pressure on the pair does not increase their concern for the group as a whole (or for the welfare of the out-group member). Instead it makes them more averse toward in-group inequality.

Column (2) characterizes social preference patterns of the out-group member. Compared to the pairs, out-group members care more about overall efficiency. Since self-interest is included in the regression, this means that the out-group member actually prefer higher payoff to the pair. They also dislike distance. Namely, the greater the difference between their own payoffs and the pairs' average payoff, the unhappier they feel. Moreover, social pressure does not significantly alter their preference patterns.

Table 4.5: In-group Measures (Pairs and Others)

VARIABLES	(1) Pairs	(2) Others
Self-Interest	0.0116	0.199*
	(0.0539)	(0.108)
Overall Efficiency	0.0916	0.304***
	(0.0649)	(0.118)
In-group Efficiency	0.320***	
	(0.0676)	
In-group Inequality	-0.210***	
	(0.0361)	
Distance		-0.327***
		(0.121)
Self-Interest × Pressure	0.0919	0.226
	(0.0878)	(0.172)
Overall Efficiency × Pressure	0.157	-0.0602
	(0.101)	(0.179)
In-group Efficiency × Pressure	-0.0590	
	(0.101)	
In-group Inequality × Pressure	-0.155**	
	(0.0615)	
Distance × Pressure		-0.141
		(0.186)
Observations	1,512	756

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

Social Preferences of Dominant lineages

A major feature of the lineage network in rural China is that the dominant lineages control the economic and political resources and influence village affairs heavily. In this section, we examine the social preference pattern of members in dominant lineages as opposed to that of members in non-dominant lineages.

The dominance of the lineage that person i belongs to is measured by variable Name, which is defined as the percentage of villagers that share the same surname as person i. Greater value of Name indicates larger lineage group, hence greater power in village affairs. The regressions we perform in this section include both main effects and Name interaction effects controlling for other personal characteristics (gender, age, education, and income). Namely, utility function is specified as

$$U_i = \sum_k \beta_k \cdot X_{ik} + \sum_k \theta_k X_{ik} \cdot Name_i + \sum_{lk} \delta_{lk} \cdot X_{ik} Z_i$$

where X_{ik} is the elements of social preference measures presented in each game (self-interest, efficiency, plus one set of inequality measure), and Z_i is the vector of personal characteristics of person i. To save space, in this section we only present the parameters of Name interaction effects, namely θ_k in different scenarios.

Columns (1) to (4) in Table 4.6 characterize Name interaction effects in Randomly Assignment. We observe that none of them is significant, indicating that when voting with randomly assigned unknown members, members of dominant lineages do not behave differently compared to members of non-dominant lineages. When it comes to Network Assignment, however, columns (5) to (8) in Table 4.6 reveal different patterns. We notice that the interaction effect for ERC is positive, indicating that members of larger lineages care less about inequality. More importantly, they seem to feel better about inequality that is to their advantage.

Table 4.6: Social Preferences of Dominant lineages (All Observations)

VARIABLES	(1) Random ERC	(2) Random FS Struct	(3) Random FS	(4) Random MM	(5) Network ERC	(6) Network FS Strict	(7) Network FS	(8) Network MM
Self-Interest × Name	-0.158	-0.145		-0.143	0.178	0.205		0.222
	(0.143)	(0.142)		(0.143)	(0.149)	(0.148)		(0.147)
Efficiency × Name	0.153	0.201	0.161	0.199	-0.289	-0.241	-0.171	-0.215
	(0.184)	(0.196)	(0.191)	(0.162)	(0.192)	(0.206)	(0.199)	(0.168)
ERC × Name	0.227				0.402*			
	(0.202)				(0.215)			
FS × Name		0.0392				0.0894		
		(0.0667)				(0.0713)		
FS Ahead × Name			-0.0226				0.313*	
			(0.155)				(0.164)	
FS Behind × Name			0.173				0.0417	
			(0.170)				(0.181)	
MM × Name				-0.112				-0.272
				(0.186)				(0.195)
Observations	2,184	2,184	2,184	2,184	2,184	2,184	2,184	2,184

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

We further break down the Name interaction effects in Network Assignment into Pairs and Others. The results are presented in Table 4.7. We notice that almost all parameters in the first 4 columns (Pairs) are statistically significant, while none of them is significant in the last 4 columns (Others). In other words, members of dominant lineage groups behave differently only when the interest of their relatives are involved. They care more about self-interest and less about the welfare of the group as a whole (Efficiency). Members of dominant lineages are also less concerned about inequality. In fact, they do not feel as bad when the inequality is to their advantage (FS Ahead). This result is consistent with phenomena reported in the literature, where larger kinship groups exploit their control of economic and political resources to suppress smaller groups.

Table 4.7: Social Preferences of Dominant lineages (Pairs and Others)

VARIABLES	(1) Pairs ERC	(2) Pairs FS Strict	(3) Pairs FS	(4) Pairs MM	(5) Others ERC	(6) Others FS Strict	(7) Others FS	(8) Others MM
Self-Interest × Name	0.398**	0.449**		0.483***	-0.446	-0.170		-0.153
	(0.174)	(0.176)		(0.179)	(0.462)	(0.407)		(0.363)
Efficiency × Name	-0.371*	-0.475*	-0.311	-0.423**	-0.0710	0.207	0.150	0.182
	(0.219)	(0.246)	(0.234)	(0.201)	(0.476)	(0.445)	(0.439)	(0.373)
ERC × Name	0.538**				0.564			
	(0.240)				(0.745)			
FS × Name		0.192**				-0.0366		
		(0.0850)				(0.180)		
FS Ahead × Name			0.670***				-0.187	
			(0.213)				(0.326)	
FS Behind × Name			0.0746				0.0402	
			(0.195)				(0.548)	
MM × Name				-0.715***				0.146
				(0.249)				(0.390)
Observations	1,442	1,442	1,442	1,442	742	742	742	742

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

Social Preferences and Personal Characteristics

Compared to the networks in western societies, the lineage ties in rural China are highly individual, in the sense that any interpersonal relationship are affected heavily by the individuals' personal characteristics. In this subsection, we examine how gender, age, education and income interact with social preference patterns.

The results are presented in Table 4.8. Overall we observe that personal characteristics interaction effects (δ_{lk}) do not differ greatly across Random Assignment and Network Assignment. Female (Gender=1) consistently display greater inequality aversion compared to male, especially when the inequality is to their disadvantage. Interestingly, female displayed greater concern for overall efficiency and the welfare of the worst-off member under Random Assignment, but such concern disappears when they are in Network Assignment. Although we do not have direct evidence to explain this phenomenon, we reason that female are by nature more caring and compassionate especially towards the deprived. However, when the interests of their relatives are involved, such nature is suppressed and outweighed by in-group concerns. In the Age panel, we observe that compared to the young, senior participants are less concerned with self-interest and inequality in both Random Assignment and Network Assignment. They seem to be less sensitive to every element in the utility function and behaves in a more neutral manner.

The interaction effects in the Education and Income panels are apparently weaker. In Network Assignment, better educated participants display greater concern for overall efficiency and tend to feel worse when the inequality is to their advantage. Such pattern, however, is not present in Random Assignment. The effect of income is almost non-existent.

Table 4.8: Social Preferences and Personal Characteristics (All Observations)

VARIABLES	(1) Random ERC	(2) Random FS Strict	(3) Random FS	(4) Random MM	(5) Network ERC	(6) Network FS Strict	(7) Network FS	(8) Network MM
Gender								
Self-Interest × Gender	0.115	0.102		0.138	0.157	0.143		0.136
	(0.0977)	(0.0953)		(0.0992)	(0.103)	(0.102)		(0.102)
Efficiency × Gender	0.0925	0.186	0.219*	0.121	-0.0397	0.0625	0.104	-0.0815
	(0.125)	(0.134)	(0.130)	(0.110)	(0.138)	(0.149)	(0.141)	(0.119)
ERC × Gender	-0.284**				-0.272*			
	(0.144)				(0.154)			
FS × Gender		-0.122***				-0.123**		
		(0.0470)				(0.0516)		
FS Ahead × Gender			-0.174				-0.146	
			(0.110)				(0.121)	
FS Behind × Gender			-0.321***				-0.348***	
			(0.114)				(0.125)	
MM × Gender				0.399***				0.221
				(0.130)				(0.136)
Age								
Self-Interest × Age	-0.0109***	-0.0104***		-0.00854**	-0.0118***	-0.0109***		-0.00923**
	(0.00371)	(0.00362)		(0.00367)	(0.00396)	(0.00383)		(0.00375)
Efficiency × Age	-0.00105	-0.00342	-0.00634	0.00240	0.00188	-0.00126	-0.00466	0.00460
	(0.00444)	(0.00471)	(0.00462)	(0.00392)	(0.00461)	(0.00492)	(0.00480)	(0.00404)
ERC × Age	0.0133**				0.0163***			
	(0.00525)				(0.00570)			
FS × Age		0.00496***				0.00608***		
		(0.00169)				(0.00182)		
FS Ahead × Age			0.00269				0.00473	
			(0.00365)				(0.00389)	
FS Behind × Age			0.0165***				0.0198***	
			(0.00455)				(0.00491)	
MM × Age				-0.00930**				-0.0144***
				(0.00464)				(0.00490)
Education								
Self-Interest × Education	0.0163	0.0133		0.0137	-0.0230	-0.0247		-0.0222
	(0.0149)	(0.0148)		(0.0150)	(0.0163)	(0.0160)		(0.0156)
Efficiency × Education	0.0183	0.00764	0.0106	0.00209	0.0376*	0.0406*	0.0310	0.0329*
	(0.0191)	(0.0200)	(0.0195)	(0.0166)	(0.0201)	(0.0211)	(0.0204)	(0.0172)
ERC × Education	-0.0286				-0.0334			
	(0.0218)				(0.0241)			
FS × Education		-0.00258				-0.0104		
		(0.00700)				(0.00761)		
FS Ahead × Education			0.00438				-0.0365**	
			(0.0155)				(0.0168)	
FS Behind × Education			-0.0128				-0.00290	
			(0.0185)				(0.0201)	
MM × Education				-0.00630				0.0246
				(0.0194)				(0.0201)
Income								
Self-Interest × Income	-2.26e-05	-2.58e-05		-8.28e-06	-3.78e-05	-3.78e-05		-3.58e-05
	(2.47e-05)	(2.15e-05)		(2.50e-05)	(2.63e-05)	(2.43e-05)		(2.63e-05)
Efficiency × Income	1.33e-05	2.07e-05	1.37e-05	2.35e-05	3.75e-05	4.84e-05	3.63e-05	3.50e-05
	(1.82e-05)	(2.23e-05)	(2.28e-05)	(1.92e-05)	(2.51e-05)	(2.96e-05)	(2.85e-05)	(2.37e-05)
ERC × Income	-2.33e-05				-5.82e-06			
	(3.56e-05)				(3.82e-05)			
FS × Income		-8.62e-06				-6.61e-06		
		(1.07e-05)				(1.27e-05)		
FS Ahead × Income			-3.49e-05*				-3.85e-05	
			(1.99e-05)				(2.50e-05)	
FS Behind × Income			-3.30e-06				1.13e-05	
			(2.92e-05)				(3.39e-05)	
MM × Income				4.28e-05				-5.51e-06
				(3.01e-05)				(3.20e-05)
Observations	2,184	2,184	2,184	2,184	2,184	2,184	2,184	2,184

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

Table 4.9 breaks down personal characteristics interaction effects in the Network Assignment into Pairs and Others. We observe that female participants in the pairs behave more selfishly compared to female participants playing the role of Others – they care more about self-interest and feel worse when the inequality is to their disadvantage. Education affects the pairs' preferences but not the Others. In particular, better educated pairs are less selfish, more averse to inequality, especially when inequality is to their advantage, and care more about the welfare of the worst-off person. Age and Income interaction effects, on the other hand, are not different across pairs and Others in an observable way.

Table 4.9: Interaction with Personal Characteristics (Pairs and Others)

VARIABLES	(1) Pairs ERC	(2) Pairs FS Strict	(3) Pairs FS	(4) Pairs MM	(5) Others ERC	(6) Others FS Strict	(7) Others FS	(8) Others MM
Gender								
Self-Interest × Gender	0.212*	0.199*		0.189	0.182	0.321		0.148
	(0.119)	(0.118)		(0.122)	(0.358)	(0.304)		(0.273)
Efficiency × Gender	-0.0177	0.0261	0.0868	-0.0965	0.0299	0.220	0.327	-0.0366
	(0.155)	(0.172)	(0.162)	(0.137)	(0.371)	(0.331)	(0.324)	(0.281)
ERC × Gender	-0.223				-0.352			
	(0.170)				(0.578)			
FS × Gender		-0.0801				-0.241*		
		(0.0593)				(0.131)		
FS Ahead × Gender			-0.0241				-0.268	
			(0.150)				(0.238)	
FS Behind × Gender			-0.300**				-0.697*	
			(0.132)				(0.404)	
MM × Gender				0.0831				0.316
				(0.171)				(0.279)
Age								
Self-Interest × Age	-0.0116**	-0.0108**		-0.0111**	-0.0175	-0.0172*		-0.0104
	(0.00464)	(0.00459)		(0.00473)	(0.0115)	(0.00917)		(0.00792)
Efficiency × Age	-2.36e-05	-0.00180	-0.00499	0.00338	0.0117	0.00510	-0.000618	0.00994
	(0.00536)	(0.00603)	(0.00584)	(0.00500)	(0.0105)	(0.00960)	(0.00936)	(0.00809)
ERC × Age	0.0156**				0.0183			
	(0.00630)				(0.0171)			
FS × Age		0.00526**				0.00801**		
		(0.00218)				(0.00392)		
FS Ahead × Age			0.00304				0.00457	
			(0.00510)				(0.00689)	
FS Behind × Age			0.0180***				0.0274**	
			(0.00544)				(0.0123)	
MM × Age				-0.0145**				-0.0140*
				(0.00636)				(0.00842)
Education								
Self-Interest × Education	-0.0321	-0.0352*		-0.0265	-0.0637	-0.0546		-0.0478
	(0.0205)	(0.0206)		(0.0211)	(0.0458)	(0.0342)		(0.0298)
Efficiency × Education	0.0264	0.0441	0.0296	0.0383	0.0547	0.0420	0.0238	0.0394
	(0.0255)	(0.0288)	(0.0273)	(0.0234)	(0.0378)	(0.0351)	(0.0338)	(0.0299)
ERC × Education	-0.0507*				0.0323			
	(0.0285)				(0.0641)			
FS × Education		-0.0213**				0.0106		
		(0.0101)				(0.0143)		
FS Ahead × Education			-0.0641**				-0.0152	
			(0.0254)				(0.0250)	
FS Behind × Education			-0.0172				0.0576	
			(0.0231)				(0.0453)	
MM × Education				0.0749**				-0.0342
				(0.0296)				(0.0312)
Income								
Self-Interest × Income	-2.99e-05	-3.20e-05		-4.70e-05	-3.57e-05	1.13e-05		-1.69e-05
	(2.71e-05)	(2.48e-05)		(2.98e-05)	(0.000113)	(8.99e-05)		(7.98e-05)
Efficiency × Income	3.35e-05	3.21e-05	2.34e-05	2.04e-05	0.000113	0.000149	0.000152	9.73e-05
	(2.46e-05)	(2.83e-05)	(2.78e-05)	(2.22e-05)	(0.000110)	(0.000101)	(9.94e-05)	(8.39e-05)
ERC × Income	-2.75e-06				-7.06e-06			
	(3.89e-05)				(0.000175)			
FS × Income		7.12e-07				-3.46e-05		
		(1.25e-05)				(4.03e-05)		
FS Ahead × Income			-2.10e-05				-6.16e-05	
			(2.42e-05)				(7.02e-05)	
FS Behind × Income			2.12e-05				-7.67e-05	
			(3.45e-05)				(0.000123)	
MM × Income				-4.33e-05				2.47e-05
				(3.42e-05)				(8.31e-05)
Observations	1,442	1,442	1,442	1,442	742	742	742	742

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

Social Preferences and Network Characteristics

The last set of results (Table 4.10) looks at how social preferences vary with network characteristics, namely how close the pairs are in terms of their relationship within the lineage network. We construct a "Relation" variable – a Coefficient of Human Relationships – to numerically express this degree of kinship in human genealogy.

The Coefficient of Human Relationships is defined over a scale of 0 to 1, where greater value indicates closer relationship. In the patriarchal lineage system in rural China, married couples form the basic units of the society, the coefficient for spouses is set to be 1. An offspring carries half the genes of each parent, so the coefficient for parent-offspring relationship is 0.5. Calculation of the coefficients for other relationships is provided in the Appendix.

Since the concept of network or relationship does not exist in Random Assignment, our analysis focuses on behaviors in Network Assignment. Moreover, since the relationship variable can only be calculated for the pairs, we exclude Others from our sample. We focus on in-group measures of social preferences in this subsection, because those are likely to be what in-group members are most concerned with and are more naturally affected by relationship. As a robustness check, we repeated the analysis using overall measures of preferences – they reveal almost identical pattern.

The upper panel of Table 4.10 presents relationship interaction effects. We observe that when voting with closer relatives, in-group members care less about self interest. We also notice that closer social relationship makes in-group members care more about both overall efficiency and in-group efficiency. In addition, in-group members playing with closer Relatives are more averse to in-group inequalities that are to their advantages. Namely, their sense of guilt increases with social proximity. The lower panel of Table 4.10 presents the triple interaction effects with relationship and with social pressure. The only statistically significant triple interaction effect is the one with in-group efficiency, meaning that pressure effect is more prominent on preference for in-group efficiency when the pair is closer. Since both overall efficiency and in-group efficiency are included in the regression, this also indicates that social pressure makes the more intimate pairs more fixated on their within-group gains even at the cost of the out-group member.

Table 4.10: Within Measure Relationship Interaction (Round 2- Pairs Only)

VARIABLES	(1) Within Strict	(2) Within
Self-Interest × Relation	-0.178*	
	(0.107)	
Overall Efficiency × Relation	0.247**	0.247**
	(0.125)	(0.125)
In-group Efficiency × Relation	0.289**	0.200*
	(0.128)	(0.117)
In-group Inequality × Relation	-0.0907	
	(0.0729)	
In-group Inequality (Ahead) × Relation		-0.180**
		(0.0878)
In-group Inequality (Behind) × Relation		-0.00180
		(0.0928)
Self-Interest × Relation × Pressure	-0.0962	
	(0.226)	
Overall Efficiency × Relation × Pressure	-0.310	-0.310
	(0.260)	(0.260)
In-group Efficiency × Relation × Pressure	0.524**	0.476*
	(0.267)	(0.244)
In-group Inequality × Relation × Pressure	0.175	
	(0.156)	
In-group Inequality (Ahead) × Relation × Pressure		0.127
		(0.184)
In-group Inequality (Behind) × Relation × Pressure		0.223
		(0.200)
Observations	1,512	1,512

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

One potential drawback of our definition of the Coefficient of Human Relationships is that although the spouses form the basis of human societies, they are in fact not genetically related. Hence as a robustness check, we separate spouses from other types of human relationships by including a dummy variable for spouses and use our former definition of the coefficient for all the other relationships. The interaction effects are reported in Table 4.11. The first two panels present the interaction terms between the utility components and spouse/relationship. Consistent with our findings in Table 4.10, participants tend to care less about self-interest and more about in-group gains when they are playing with a spouse or a relative with closer kinship. Interestingly, when participants vote with a spouse, they tend to be more averse to in-group inequalities that are to their advantages. However when they play with a close kin, they tend to be more averse to in-group inequalities that are to their disadvantages. The last two panels show the triple interaction effects among utility components, spouse/relationship, and the social pressure treatment dummy. We find that the triple interactions with spouse dummies are not statistically significant. The only significant triple interaction with the coefficient of human relationships is in-group inequality, in particular when the in-group inequality is to one's disadvantage. Therefore, the intensified aversion toward in-group inequality under social pressure we observed in Table 4.10 is mainly driven by degree of kinship outside marriage.

Table 4.11: Within Measure Spouse/Relation Interaction (Round 2- Pairs Only)

VARIABLES	(1) Within Strict	(2) Within
Self-Interest × Spouse	-0.291**	
	(0.130)	
Overall Efficiency × Spouse	0.121	0.121
	(0.151)	(0.151)
In-group Efficiency × Spouse	0.298*	0.152
	(0.154)	(0.142)
In-group Inequality × Spouse	-0.0815	
	(0.0882)	
In-group Inequality (Ahead) × Spouse		-0.227**
		(0.104)
In-group Inequality (Behind) × Spouse		0.0639
		(0.114)
Self-Interest × Relation	-1.056**	
	(0.464)	
Overall Efficiency × Relation	-0.593	-0.593
	(0.569)	(0.569)
In-group Efficiency × Relation	1.131*	0.603
	(0.606)	(0.550)
In-group Inequality × Relation	0.114	
	(0.305)	
In-group Inequality (Ahead) × Relation		-0.414
		(0.386)
In-group Inequality (Behind)× Relation		0.642*
		(0.381)
Self-Interest × Spouse × Pressure	0.0401	
	(0.131)	
Overall Efficiency × Spouse × Pressure	0.108	0.108
	(0.154)	(0.154)
In-group Efficiency × Spouse × Pressure	0.226	0.246
	(0.165)	(0.152)
In-group Inequality × Spouse × Pressure	-0.0725	
	(0.0908)	
In-group Inequality (Ahead) × Spouse × Pressure		-0.0524
		(0.111)
In-group Inequality (Behind)× Spouse × Pressure		-0.0925
		(0.113)
Self-Interest × Relation × Pressure	0.577	
	(0.483)	
Overall Efficiency × Relation × Pressure	0.776	0.776
	(0.550)	(0.550)
In-group Efficiency × Relation × Pressure	-0.679	-0.391
	(0.562)	(0.504)
In-group Inequality × Relation × Pressure	-0.919***	
	(0.348)	
In-group Inequality (Ahead) × Relation × Pressure		-0.631
		(0.417)
In-group Inequality (Behind)× Relation × Pressure		-1.208***
		(0.430)
Observations	1,512	1,512

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

4.5 Conclusion

Economists in the past two decades have come to realize that human choices are not only driven by self-interest but also social preferences a persons concern over resources allocated to other people. Moreover, such preferences may be affected by the environment in which such choices are made, especially social networks and social pressure. To overcome the difficulties of identifying causal effects in observational studies and the lack of interpersonal relationships in lab experiments, we performed a lab-in-the-field experiment in rural China, where the lineage network structure is salient and relevant, and invited villagers to bring in their real-world social contacts to play an array of three-person allocation games. The variations in payoffs and the randomization of social pressure treatment allows us to identify social preference patterns as well as network effects and pressure effects. Our post-experiment survey collects information on demographics and the specifics of the interpersonal relationship between each pair of participants, which allows us to investigate (1) how participants perceive resources allocated to in-group and out-group members; (2) if members of dominant social groups have different social preference patterns; (3) if social preferences vary with personal characteristics such as gender and age; (4) if variations in social preferences can be explained by social proximity.

Consistent with evidence shown in prior literature, we find that (a) in addition to self-interest, participants consistently display preference for social efficiency and distaste for inequality. Such social preferences are intrinsic and robust against alternative specifications. Moreover, social network or social pressure alone cannot alter behaviors. Social preference patterns, in particular inequality aversion, change only when the two forces are combined; (b) In-group members tend to be fixated on in-group gains and disregard the welfare of out-group members, and such fixation is often exacerbated by social pressure; (c) When playing with people in their networks, members of larger lineage groups, especially those in the position of pairs, tend to be less concerned with overall efficiency and inequality. They also seem not to feel as bad when the inequality is to their advantage; (d) Social preference patterns also vary with personal characteristics, such as gender, age, and education; (e) Closer social relationship makes in-group members care more about efficiency and inequality. Moreover, social pressure makes the more intimate pairs more fixated on within-group gains even at the cost of the out-group member.

This study is to our knowledge the first lab-in-the-field experiment that embeds both real-world social network and social pressure into allocation games. Our results not only provide empirical evidence for the social preference theories but also offer policy insights for the developing world. The democratization progress of many developing countries in Asia and Africa has been complicated by social network structures, especially when the in-group out-group division is salient. In this case, recognizing the possibility that in-group preferences may outweigh out-group preferences and yield economically inefficient outcomes is essential. Since social pressure and social network tend to have a combined effect that reinforces each other, policy makers must be careful in choosing an appropriate voting method. Secret ballot instead of show of hands, for example, may help alleviate in-group members' fixation on in-group gains when they are under social pressure, and may help improve the welfare of the out-group members. In addition, since members of the dominant lineages consistently displayed less pro-social tendencies during the experiment, constraining the power they can wield in the decision-making process is important in enhancing the performance of rural democracy projects. Last but not least, since better educated participants are more pro-social during our experiment – they care more about overall efficiency and the welfare of the worst-off member and are more averse to inequalities that are to their advantages – having them more involved in village decision making could be beneficial.

References:

[1] Akerlof, G. A. (1997). Social distance and social decisions. Econometrica: Journal of the Econometric Society, 1005-1027.

[2] Alesina, A., Baqir, R., & Easterly, W. (1997). Public goods and ethnic divisions (No. w6009). National Bureau of Economic Research.

[3] Andreoni, J. (1988). Privately provided public goods in a large economy: the limits of altruism. Journal of Public Economics, 35(1), 57-73.

[4] Andreoni, J. (1989). Giving with impure altruism: applications to charity and Ricardian equivalence. The Journal of Political Economy, 1447-1458.

[5] Bandiera, O., Barankay, I., & Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. The Quarterly Journal of Economics, 120(3), 917-962.

[6] Banerjee, A., Iyer, L., & Somanathan, R. (2005). History, social divisions, and public goods in rural India. Journal of the European Economic Association, 3(2-3), 639-647.

[7] Baker, H. D. (1979). Chinese family and kinship. New York: Columbia University Press

[8] Bernhard, H., Fehr, E., & Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. American Economic Review, 96(2), 217-221.

[9] Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. American economic review, 166-193.

[10] Bolton, G. E., & Ockenfels, A. (2006). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments: comment. American Economic Review, 96(5), 1906-1911.

[11] Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. Econometrica, 82(4), 1273-1301.

[12] Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. Quarterly journal of Economics, 817-869.

[13] Chen, Y., & Li, S. X. (2009). Group identity and social preferences. American Economic Review, 99(1), 431-57.

[14] Coate, S., & Ravallion, M. (1993). Reciprocity without commitment: Characterization and performance of informal insurance arrangements. Journal of development Economics, 40(1), 1-24.

[15] Coleman, J. S. (1988). Social capital in the creation of human capital. American journal of sociology, S95-S120.

[16] Cohen, M. L. (2005). Kinship, contract, community, and state: Anthropological perspectives on China. Stanford University Press.

[17] Cox, J. C., Sadiraj, K., & Sadiraj, V. (2001). Trust, Fear, Reciprocity, and Altruism. University of Arizona, mimeo.

[18] Engelmann, D., & Strobel, M. (2004). Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. The American Economic Review, 94(4), 857-869.

[19] Fehr, E., & Gchter, S. (1999). Cooperation and punishment in public goods experiments. Institute for Empirical Research in Economics working paper, (10).

[20] Fehr, E., & Gchter, S. (2002). Altruistic punishment in humans. Nature, 415(6868), 137-140.

[21] Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. Quarterly journal of Economics, 817-868.

[22] Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruismexperimental evidence and new theories. Handbook of the economics of giving, altruism and reciprocity, 1, 615-691.

[23] Fukuyama, F. (2001). Social capital, civil society and development. Third world quarterly, 22(1), 7-20.

[24] Guala, F. & Filippin, A. (2017), The Effect of Group Identity on Distributive Choice: Social Preference or Heuristic?. Econ J, 127: 1047-1068. doi:10.1111/ecoj.12311

[25] Gchter, S., & Fehr, E. (1999). Collective action as a social exchange. Journal of Economic Behavior & Organization, 39(4), 341-369.

[26] Guth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. Journal of economic behavior & organization, 3(4), 367-388.

[27] Halali, E., Kogut, T., & Ritov, I. (2017). Reciprocating (more) specifically to you: The role of benefactor's identifiability on direct and upstream reciprocity. Journal of Behavioral Decision Making, 30(2), 473-483.

[28] He, Q., Pan, Y., & Sarangi, S. (2017). Lineage-based heterogeneity and cooperative behavior in rural China. Journal of Comparative Economics.

[29] He, X., (2003). New Rural China. Nanning: Guangxi Normal University Press

[30] Hu, B. (2007). Informal institutions and rural development in China. London: Routledge.

[31] Huang, P. C. (2008). Centralized minimalism: semiformal governance by quasi officials and dispute resolution in China. Modern China, 34(1), 9-35.

[32] Kollock, P. (1998). Transforming social dilemmas: group identity and co-operation. Modeling rationality, morality, and evolution, 7, 185-209.

[33] Landa, J. T. (1997). Trust, ethnicity, and identity: beyond the new institutional economics of ethnic trading networks, contract law, and gift-exchange. University of Michigan Press.

[34] Liu, L., & Murphy, R. (2006). Lineage networks, land conflicts and rural migration in late socialist China. The Journal of Peasant Studies, 33(4), 612-645.

[35] Lu, Y., & Tao, R. (2017). Organizational Structure and Collective Action: Lineage Networks, Semiautonomous Civic Associations, and Collective Resistance in Rural China. American Journal of Sociology, 122(6), 1726-1774.

[36] Masclet, D., Noussair, C., Tucker, S., & Villeval, M. C. (2003). Monetary and nonmonetary punishment in the voluntary contributions mechanism. American Economic Review, 93(1), 366-380.

[37] Messer, K. D., Poe, G. L., Rondeau, D., Schulze, W. D., & Vossler, C.

A. (2010). Social preferences and voting: An exploration using a novel preference revealing mechanism. Journal of Public Economics, 94(3-4), 308-317.

[38] Miquel, G. P., Qian, N., & Yao, Y. (2012) Social Fragmentation, Public Goods and Elections: Evidence from China, NBER Working Paper No. 18633

[39] Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. Journal of Public Economics, 89(11-12), 2325-2368.

[40] Nee, V., & Ingram, P. (1998). Embeddedness and beyond: institutions, exchange, and social structure. The new institutionalism in sociology, 19.

[41] Nickerson, D. W. (2008). Is voting contagious? Evidence from two field experiments. American Political Science Review, 102(01), 49-57.

[42] Postlewaite, A. (1998). The social basis of interdependent preferences. European Economic Review, 42(3), 779-800.

[43] Pan, Y. (2011). Born with the Right Surname? Lineage Networks and Political and Economic Opportunities in Rural China. Louisiana State University, Working Paper, (15).

[44] Peng, Y. (2004). Kinship networks and entrepreneurs in Chinas transitional economy. American Journal of Sociology, 109(5), 1045-1074.

[45] Peng, Y. (2010). When formal laws and informal norms collide: Lineage networks versus birth control policy in China. American Journal of Sociology, 116(3), 770-805.

[46] Rabin, M. (1993). Incorporating fairness into game theory and economics. The American economic review, 1281-1302.

[47] Rege, M., & Telle, K. (2004). The impact of social approval and framing on cooperation in public good situations. Journal of public Economics, 88(7), 1625-1644.

[48] Thurston, A. F. (1998). Muddling toward Democracy: political change in grassroots China (Vol. 31, No. 19-23). United States Institute of Peace.

[49] Tiebout, C. M. (1956). A pure theory of local expenditures. The journal of political economy, 416-424.

[50] Tsai, L. L. (2007). Accountability without democracy: Solidary groups and public goods provision in rural China. Cambridge University Press.

[51] Van Dijk, F., Sonnemans, J., & van Winden, F. (2002). Social ties in a public good experiment. Journal of Public Economics, 85(2), 275-299.

[52] Xiao, T., (2001). Lineage in Village Governance: Research Based on Nine Villages. Shanghai: Press of Shanghai Book Store

[53] Zimmerman, J. F. (1988). Referendum Voting: Social Status and Policy Preferences. By Hahn Harlan and Kamieniecki Sheldon. Westport, CT: Greenwood, 1987. 150p. $35.00. American Political Science Review, 82(04), 1377-1378.

CHAPTER 5

CONCLUSION

The above three chapters investigate preferences with respect to different aspects of human life and choices made at various levels of human society. The first chapter examines preferences at the individual level and focuses on occupational choices of migrant workers and workers in China. The second chapter looks at household-level decisions and how that could affect human capital formation of school-age children in Kyrgyz Republic. The third chapter goes beyond individuals and households to examine how social preferences and social network can affect collective decisions at the community level in this case where Chinese villagers vote for public goods provision.

The key message I wish to convey through this collection of essays is that in all areas of human life and at all levels of the human society, preferences and choices are not only the causes but also the consequences of inequality.

First of all, preferences of and choices made by individuals, households, and communities have profound implications for inequalities in multiple aspects of human society. In the first chapter of this dissertation, we used a labor search model with market frictions to theoretically demonstrate that the heterogeneity in preferences across two groups migrant workers and urban workers may negate compensating differentials and lead to the counterintuitive phenomenon in the Chinese labor market where migrant workers take blue-collar jobs with undesirable working conditions and are nonetheless paid less. In the second chapter, we find that households preferences over the allocation of financial and labor resources, especially their emphasize on short-term consumptions and use of child labor to compensate for adult labor insufficiency, may have led to the decrease in educational expenditure and the deterioration of school attendance, placing school-age children within migrant households in a disadvantaged position. In the final chapter, we showcased that social preferences a persons concern over the welfare of other people may be conditioned by membership within a social network and exposure to social pressure. In particular, our finding that members of the dominant lineages tend to behave less pro-socially in thee social context helps explain why the welfare of underrepresented lineages in rural China is overlooked.

However, it is not our intention to blame individual or collective choices for all the dire manifestations of inequalities in our society, since the preferences of each person, each household, and each community are shaped by the evolution of the human society itself and are therefore consequences of inequality as well. For example, in the first chapter, we demonstrated through a choice experiment that preferences of the migrant workers are more dispersedly distributed and vary more with personal characteristics such as gender and age. These are evidence that preferences of the individuals are not exogenously imposed but are informed by the role that they play in the society. In the second chapter, although we found that households in the Kyrgyz Republic would rather spend remittances on durable goods than invest them in education, we reason that such preferences are plausibly informed by the low expected return to schooling due to the low quality of post-secondary education. Similarly, these households decisions to retain school-age children at home for house work in order to compensate for the insufficiency of adult labor force are likely to be a consequence of the underperformance of domestic labor market, which caused large waves of labor out-migration in the first place. Last but not least, in the final chapter, although villagers preferences for in-group gains at the cost of out-group members welfare contribute to the inequality in rural China as dominant lineages manipulate village elections and public goods voting, the formation of such preferences is an endogenous process and is affected by inequality itself. For example, we find that social preference patterns vary with personal characteristics, with more educated villagers caring more about the overall efficiency of resource allocation. The underinvestment of schooling and low level of education in rural China therefore at least partly contributes to the lack of prosocial behaviors in our study.

Informed policy interventions that take the heterogeneity of preferences into account are therefore necessary to rectify uninformed decisions and alleviate inequality in the developing world. In the first chapter, for example, we argue that although the implementation of labor contract law and workplace safety regulations is well intended, it may in fact backfire when the two groups on the labor market the migrant workers and the urban workers have heterogeneous preferences as the new policies may induce urban workers into the market and crowd out migrant workers. In the second chapter, our finding that remittance-receiving households prefer short-term expenditures (durable goods) over long-term investments (education) and prefer to retain children in farm work calls for the need for monitoring farm labor hours of school-age children as well as the provision of financial literacy programs that help household heads balance short-term and long-term financial goals. Moreover, targeted investment to improve the quality of education services in the country may help increase perceived return to schooling a first step to alter preference patterns at the household level. In the final chapter, our study not only provides empirical evidence for the social preference theories but also urges policy makers to be careful in choosing an appropriate voting method for example secret ballot instead of show of hands. In addition, constraining the power of dominant lineage and having better educated villagers more involved in village affairs could be welfare improving and inequality alleviating.

APPENDIX A

APPENDIX FOR CHAPTER 2

Experiment Materials

Suppose you are unemployed at the moment and are actively looking for a job. You will be presented with 4 hypothetical choice scenarios in the following questionnaire. Each choice scenario consists of two hypothetical jobs that you are offered – they are the only offers you have. You can compare the two options and choose one, or reject both and stay unemployed for another 6 months. Please note that while you stay unemployed, you will not receive any wage payment or benefits (insurance, unemployment benefits, housing, etc.). Here are some detailed explanations for some terms used in the choice scenarios:

Environment: "indoor but not office" refers to factories, shopping malls, restaurants, hotels, or residentials; "outdoor" refers to construction sites, train stations, or roads;

Contract: The formal employment contract signed between the employer and the employee;

Insurance: Health/ Unemployment/ Injury/ Childcare Insurance, Pension, and Housing Fund;

Danger: Serious threat to health and safety (risk of accidents, exposure to toxics, and chronic damages from working night shifts, etc.);

Location: "First-line" refers to Beijing, Shanghai, Guangzhou etc.; "Second-line" refers to Jinan, Harbin, etc.;

Please put a check mark " √ " on the row labeled "Your Choice".

<u>Version 1</u>

V1E1、 （Scenario 1）

	Alternative 1	**Alternative 2**	**Alternative 3 (Stay Home)**
Wage	3000	5000	
Hours	8 hours/day 5 days/week	8 hours/day 5 days/week	
Environment	Indoor, but not office	Outdoor	NA
Contract	Yes	No	
Insurance	Yes	Yes	
Danger	High	Low	
Location	Second-line	Second-line	Current City/ Hometown
Your Choice			

V1E2、 （Scenario 2）

	Alternative 1	**Alternative 2**	**Alternative 3 (Stay Home)**
Wage	5000	4000	
Hours	8 hours/day 5 days/week	8 hours/day 5 days/week	
Environment	Office	Office	NA
Contract	No	Yes	
Insurance	No	No	
Danger	Low	Low	
Location	Second-line	Second-line	Current City/ Hometown
Your Choice			

V1E3、 （Scenario 3）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	3000	5000	
Hours	11 hours/day 7 days/week	11 hours/day 7 days/week	
Environment	Outdoor	Outdoor	NA
Contract	No	Yes	
Insurance	No	Yes	
Danger	Low	High	
Location	Second-line	Town	Current City/ Hometown
Your Choice			

V1E4、 （Scenario 4）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	5000	5000	
Hours	9 hours/day 6 days/week	8 hours/day 5 days/week	
Environment	Office	Outdoor	NA
Contract	Yes	No	
Insurance	No	No	
Danger	Middle	Middle	
Location	Town	First-line	Current City/ Hometown
Your Choice			

Version 2

V2E1、 （Scenario 1）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	4000	5000	NA
Hours	10 hours/day 7 days/week	9 hours/day 6 days/week	
Environment	Outdoor	Outdoor	
Contract	No	No	
Insurance	No	Yes	
Danger	Low	Low	
Location	First-line	Second-line	Current City/ Hometown
Your Choice			

V2E2、 （Scenario 2）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	3000	4000	NA
Hours	10 hours/day 7 days/week	11 hours/day 7 days/week	
Environment	Outdoor	Office	
Contract	Yes	Yes	
Insurance	Yes	Yes	
Danger	High	High	
Location	First-line	Second-line	Current City/ Hometown
Your Choice			

V2E3、 （Scenario 3）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	6000	4000	
Hours	10 hours/day 7 days/week	10 hours/day 7 days/week	
Environment	Outdoor	Outdoor	NA
Contract	No	Yes	
Insurance	No	Yes	
Danger	Middle	Middle	
Location	First-line	Town	Current City/ Hometown
Your Choice			

V2E4、 （Scenario 4）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	3000	3000	
Hours	10 hours/day 7 days/week	10 hours/day 7 days/week	
Environment	Office	Outdoor	NA
Contract	No	Yes	
Insurance	Yes	No	
Danger	Middle	High	
Location	Town	Town	Current City/ Hometown
Your Choice			

Version 3

V3E1、 （Scenario 1）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	3000	4000	NA
Hours	9 hours/day 6 days/week	11 hours/day 7 days/week	
Environment	Indoor, but not office	Outdoor	
Contract	Yes	Yes	
Insurance	No	No	
Danger	Middle	High	
Location	First-line	Second-line	Current City/ Hometown
Your Choice			

V3E2、 （Scenario 2）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	5000	3000	NA
Hours	11 hours/day 7 days/week	10 hours/day 7 days/week	
Environment	Office	Indoor, but not office	
Contract	No	Yes	
Insurance	Yes	Yes	
Danger	High	Middle	
Location	Town	Town	Current City/ Hometown
Your Choice			

V3E3、 （Scenario 3）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	3000	3000	NA
Hours	10 hours/day 7 days/week	8 hours/day 5 days/week	
Environment	Indoor, but not office	Outdoor	
Contract	Yes	Yes	
Insurance	No	Yes	
Danger	Low	Middle	
Location	Town	Second-line	Current City/ Hometown
Your Choice			

V3E4、 （Scenario 4）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	3000	3000	NA
Hours	10 hours/day 7 days/week	8 hours/day 5 days/week	
Environment	Indoor, but not office	Indoor, but not office	
Contract	Yes	No	
Insurance	No	Yes	
Danger	High	Low	
Location	Second-line	Town	Current City/ Hometown
Your Choice			

Version 4

V4E1、 （Scenario 1）

	Alternative 1	**Alternative 2**	**Alternative 3 (Stay Home)**
Wage	5000	3000	NA
Hours	8 hours/day 5 days/week	9 hours/day 6 days/week	
Environment	Outdoor	Office	
Contract	Yes	Yes	
Insurance	No	No	
Danger	Middle	Middle	
Location	First-line	Town	Current City/ Hometown
Your Choice			

V4E2、 （Scenario 2）

	Alternative 1	**Alternative 2**	**Alternative 3 (Stay Home)**
Wage	5000	5000	NA
Hours	10 hours/day 7 days/week	9 hours/day 6 days/week	
Environment	Indoor, but not office	Indoor, but not office	
Contract	No	No	
Insurance	Yes	Yes	
Danger	Middle	High	
Location	Second-line	Town	Current City/ Hometown
Your Choice			

V4E3、 （Scenario 3）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	4000	3000	
Hours	8 hours/day 5 days/week	11 hours/day 7 days/week	
Environment	Office	Office	NA
Contract	Yes	No	
Insurance	Yes	Yes	
Danger	Middle	Middle	
Location	Second-line	Second-line	Current City/ Hometown
Your Choice			

V4E4、 （Scenario 4）

	Alternative 1	Alternative 2	Alternative 3 (Stay Home)
Wage	6000	3000	
Hours	11 hours/day 7 days/week	10 hours/day 7 days/week	
Environment	Office	Office	NA
Contract	No	Yes	
Insurance	No	Yes	
Danger	Middle	Low	
Location	First-line	First-line	Current City/ Hometown
Your Choice			

Cornell University/ Shandong University of Finance and Economics

Questionnaire About Job Choices by Migrant and Urban Workers

A. Basic Information

A1. Your gender: Male Female

A2. Your age: ___________

A3. Highest level of education achieved: ：
①None ②Elementary ③Junior High ④High School ⑤2-year Professional ⑥Bachelor's ⑦ Master's

A4. Hukou location: ________ Province ________City________Town________Village

A5. Hukou type: ①Rural ②Urban

A6. Marital status: ①Married ②Single/ Divorced/ Widowed

A7. How many children do you have? ________

A8. In the past year (2016) what is the total monthly income (including yourself, your spouse, and children) of your household? ________Yuan RMB

B. Work Situation

(B1-B3 are for migrant workers only)

B1. Have you worked in the city as a migrant worker in the past year (2016)? ①Yes ②No

B2. How many months did you work as a migrant worker in the city in 2016? ________

B3. How many days per month (on average) did you work in the city in 2016? ________

B4. The industry you work in:
① Manufacture ② Construction ③Retail ④Residential Services/ Maintenance
⑤ Hotel and Catering ⑥Transportation/ Storage/ Postal Service ⑦ Other: ________

B5. Your current occupation: ________

B6. Which year did you start your current occupation? ________

B7. On average, I work ________ hours a day, ________ days a week.

B8. What is your monthly income? ________ Yuan RMB
B8_1: Including wage ________ Yuan RMB;
B8_2: Bonus and benefits ________ Yuan RMB;
B8_3: Other payments (in-kind transfer, etc.) ________ Yuan RMB;

B9. Please choose if the following items properly describe your working environment:

	Yes	No
B9_1. I work as a laborer (heavy physical activities)		
B9_2. I work outdoor		
B9_3. My job involves serious danger (explosion, falling, etc.)		
B9_4. My job threatens my health (exposure to chemicals and toxics, etc.)		
B9_5. I have to endure heat in my job		
B9_6. I have to work night shifts		
B9_7. I have to stand most of the time		
B9_8. I have to endure bad smell and dirty substances		

B10. Do you have unemployment insurance now?
① Paid by employer ② Paid by myself ③ Jointly paid by employer/self ④ No

B11. Are you enrolled in pension plan now?
① Paid by employer ② Paid by myself ③ Jointly paid by employer/self ④ No

B12. Do you have injury insurance now?
① Paid by employer ② Paid by myself ③ Jointly paid by employer/self ④ No

B13. Are you enrolled in housing fund now?
① Paid by employer ② Paid by myself ③ Jointly paid by employer/self ④ No

B14. Have you signed contract with your employer?
① Yes ② No
B14_1. If yes, what is the nature of the contract?
① Permanent ② Long-term (1 year or above) ③ Short-term (below 1 year)

APPENDIX B

APPENDIX FOR CHAPTER 3

Table B.1: Kyrgyz: Reliance on Remittances

	Overall	**Receiving**
2010	8.51%	84.3%
2011	10.12%	83.2%
2012	10.44%	79.8%
2013	8.97%	63.6%
2016	7.98 %	75.1%

Table B.2: Kyrgyz: Percentage Spent on Education and Other Items

	2010	2011	2012	2013	2016
Education	5.53%	6.17%	4.55%	4.57%	5.73%
Food	49.93%	55.82%	43.58%	43.61%	48.53%
Non-durable	13.79%	2.96%	18.73%	15.88%	17.64%
Durable	1.63%	1.82%	4.06%	5.41%	4.54%
Wedding	5.76%	4.95%	5.05%	6.79%	3.45%
Utilities	11.78%	14.97%	11.99%	11.32%	8.99%
Health	1.83%	2.45%	2.37%	2.14%	2.54%
Other	9.75%	10.86%	9.66%	10.88%	8.58%

Table B.3: Kyrgyz: Remittance Trend (By Urban/Rural)

	2010	2011	2012	2013	2016
Urban					
Number of Migrants Per Household	0.134	0.143	0.154	0.155	0.0877
Received Remittances	7.28%	8.65%	8.83%	7.07%	5.95%
Total Annual Amount of Cash Remittances	6,826	13,917	16,489	12,695	7,288
Total Annual Amount of In-Kind Remittances	37.6	107	73.43	71.72	26.1
Per Capita Annual Remittances	1,333	2,491	3,520	2,300	1,342
Rural					
Number of Migrants Per Household	0.207	0.22	0.289	0.304	0.234
Received Remittances	11.90%	14.50%	15.70%	18.20%	13.20%
Total Annual Amount of Cash Remittances	8,130	31,283	26,504	18,997	25,034
Total Annual Amount of In-Kind Remittances	22.29	73.91	158.7	136	119.4
Per Capita Annual Remittances	1,482	5,436	5,054	3,077	3,689

Table B.4: Education Expenditure Comparison: Receiving and Non-Receiving Families

	2010		2011		2012		2013		2016	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Total Houshold Education Expenditure	2602	2417	3043	2775	3371	3281	4281	3909	5118	4487
Tuition	587.1	487.6	580.1	310.0	582.9	405.7	652.2	337.6	716.0	335.2
School Supplies	1825	1746	2238	2195	2536	2506	3310	3367	3791	3237
Other Education Expenses	190.5	183.3	224.4	270.3	252.7	369.1	238.0	173.4	327.9	264.0
Tutoring	0	0	0	0	0	0	80.61	30.48	166.9	138.9

Table B.5: Reasons Not To Enroll (Primary)

Reasons	Percent
Costs too much	1.44
School is too far	0.66
Illness	0.66
Does not like study	0.26
Conflict with Pupils or Teacher	0.13
Will start next year	93.05
Finished	0.26
Other	3.54

Table B.6: Reasons Not To Enroll (Secondary)

Reasons	Percent
Costs too much	4.75
School is too far	3.06
Illness	3.74
Does not like study	15.11
Works to support family	9.17
Will start next year	1.02
Finished	53.31
Political unrest	0.51
Other	9.34
Observatrions	589

Table B.7: Use of Remittances (Ranked High to Low)

	2010	2011	2012	2013	2016	Overall
Non-durable Goods	71.7%	72.0%	70.0%	66.6%	76.1%	71.0%
Saving	23.0%	41.5%	44.6%	32.3%	42.4%	36.9%
Wedding	26.7%	30.8%	26.7%	26.7%	13.3%	25.4%
Education	20.7%	22.5%	28.7%	23.3%	30.3%	25.0%
Durable Goods	18.3%	26.2%	29.8%	17.7%	24.6%	23.4%
Health and Medication	19.7%	17.0%	25.1%	22.8%	29.5%	22.6%
Other	17.0%	11.5%	7.7%	7.9%	12.9%	11.1%
Funeral	6.7%	9.8%	12.7%	6.2%	11.4%	9.3%
Gifts	4.0%	6.1%	3.9%	3.4%	13.6%	5.8%

Table B.8: Summary Statistics for Regression Variables

Variable	Mean	SD	Min	Max
Dependent Variables				
Household Education Expenditure	9,112	12,558	0	278,900
Non-Durable Goods Expenditure	23,066	25,949	0	544,000
Wedding Expenditure	9,528	23,743	0	500,000
Utilities Expenditure	21,000	52,985	0	4,812,000
Other Expenditure	18,346	21,688	0	345,000
Food Expenditure	76,154	44,406	0	893,365
Durable Goods Expenditure	7,296	24,984	0	1,275,000
Healthcare Expenditure	4,701	9,869	0	342,000
School Attendance Rate	0.564	0.496	0	1
Independent Variables				
Total Household Remittances	17,789	145,774	0	9,876,000
Per Capita Remittances	3,124	27,748	0	1,411,000
Age of Household Head	52	14	16	105
Percentage of Male Household Heads	78.8%			
Household Size	4.9	2.4	1.0	17.0
Household Head Education Level	4.7	1.5	1.0	8.0
Percentage of Household Heads Married	71.0%			
House Value	1,185,000	1,488,000	6,000	14,000,000
Age of School-age Children	15.45	5.544	6	24
Percentage of Female School-age Children	50.7%			
Hours per Day Homework	1.759	1	0	20
Hours per Day House/Farm Work	1.37	1.392	0	20
Hours per Day Outside Work	0.0572	0.481	0	10

Table B.9: First Stage Results

	Educ Expense 1	**Educ Expense 2**	**Other Expenses**	**Attendance**
	Total Remittances	PC Remittances	PC Remittances	PC Remittances
Instrument				
Drought * Land Area	0.00246**	0.00204**		
	-0.0009	-0.0008		
Distance to Road			0.0000962***	
			0	
Average Distance				0.0000614***
				0
Covariates				
Age of Household Leader	0.00532	0.00134	0.0053**	0.0171**
	-0.0192	-0.016	-0.0074	-0.0131
Gender of Household Leader	-2.328**	-1.940**	-0.434***	-0.552***
	-0.757	-0.6282	-0.2428	-0.2596
Household Size	0.0409	0.0195	0.1301*	0.0621*
	-0.0951	-0.0789	-0.0321	-0.027
Education Level of Household Leader	-0.119	-0.0874	-0.0008	-0.00574
	-0.1444	-0.1198	-0.0668	-0.0486
Marital Status of Household Leader	0.672	0.563	0.183	0.159
	-0.5714	-0.4741	-0.1951	-0.1696
House Value	0.000000241**	0.000000206**	0	0
	0	0	0	0
Age of Child				0.0651***
				-0.0132
Gender of Child				-0.326
				-0.2597
Ethnicity Control	Yes	Yes	Yes	Yes
Cragg-Donald F-Statistics	6.96	6.96	13.83	14.53
P-Value	0.0083	0.0084	0.0002	0

Table B.10: Itemized Household Expenditure

Category	Items included
Food	Food
Non-durable Goods	Clothing; Shoes; Soap, detergents; Personal care items and cosmetics
Durable Goods	Cell and stationary phone; Furniture and other interior equipments; Electronics and spare parts; Electric and household appliances; Other durable goods
Wedding	Celebrations, funerals, rituals
Utilities	Electricity; Cold water and sewage; Hot water; Central heating; Gas (natural and liquified); Coal and other fuel for heating; Construction and maintenance/repair of housing; Maintenance and repair of household vehicles and appliances
Health	Medicines; Medical care, including dental care
Other	Transportation services; All types of taxes (income, land, etc) and social benefit plan contributions; Entertainment, recreation, eating out; Internet

APPENDIX C

APPENDIX FOR CHAPTER 4

Part A: Experiment Materials

A1. Recruitment Material

Cornell University/Shandong University of Finance and Economics

Survey on Rural Social Networks and Preference for Public Goods

Participants Recruitment Material

Study Name	Voting Games – A Social Network Study
Study Type	**Lab-in-the-Field Experiment** After obtaining village leader's consent, we ask the village leaders to send out notices to households in the village and ask if they would like to participate in an experiment to earn some cash (50-90 RMB in total depending on their performance). If they agree, we will read them the oral consent and ask them to go to the specified location at the designated time.
Pay	All participants will earn cash payments ranging from 50RMB to 90RMB, depending on performance in the experiment
Duration	45- 60 minutes
Abstract	Play as voters in a series of allocation games
Description	To participants: In this experiment you and the other 8 participants will be asked to vote on a series of allocation plans. The experiment will take at most 45 minutes and will be followed by a short survey (where we ask for some basic information about yourself and your social interactions) that will take 15 minutes at most. Please feel free to bring along a relative or a friend of your own choice to the experiment. You can also come alone, and this will not affect your potential earning. Throughout the experiment you will get a compensation of 50 – 90 RMB, depending on your performance. The information you provide in the experiment as well as the survey will not be identifiable and will remain confidential.
Eligibility Requirements	Permanent resident of the sampled village. 18 to 65years old. A balanced sample of male and female. Mentally and physically healthy.
Experimenter	Xin Gao (PhD student at Cornell University), xg68@cornell.edu
Deadlines	Sign-Up: 1 hour(s) before the appointment Cancellation: 24 hour(s) before the appointment

A2. Experiment Instruction

Rural Social Network and Preferences in Voting

Experiment Instructions

First of all, we would like to welcome you to today's experiment. The experiment is a research project of our university that aims to understand how villagers make choices in voting.

Today's experiment has **2 Rounds**, with **1 Practice Game and 6 Games** in each round. The experiment will take anywhere between 30 and 45 minutes. In each game, you will be assigned to a 3-person group along with two other people in this room to vote on a payoff allocation plan. The voting outcome of your group will determine your final monetary payoff in this experiment, so we encourage you to consider your decisions carefully. Please do not talk to other participants during the experiment. Please raise your hand and let us know if you have questions.

Round 1

We now begin **Round 1** of the experiment. In this round, we will randomly assign any 3 people in this room to form a group. You will not know who your group members are, and your group members will not know who you are. In each game, you will receive 2 "plates": a Red Plate and a Blue Plate. You will see yellow labels indicating ownership – who will get what amount of money if a certain plate is chosen – on the plates. For example, the row that is labelled "You" shows your payoffs if Red or Blue is chosen. The two rows labelled "Other" show payoffs to the other two members in your group if Red or Blue is chosen. We have also included the numerical values of payoffs in the brackets underneath the yellow labels.

The rules are as follows. In each game, the 3-person group will vote on an allocation plan by choosing either Red or Blue. The voting game follows a majority rule. Namely, if two or more than two members in a 3-person group voted for Red, then for this group we will implement the allocation plan shown on the Red Plate. That is, every member in this 3-person group will get the amount labelled "you" on the Red Plate. Similarly, if two or more members in a 3-person group voted for Blue, then for this group we will implement the Blue Plate. That is, every member in this 3-person group will get the amount labelled "you" on the Blue Plate. After each game, the experimenters will calculate votes and distribute payoff in cash to each participant.

Read to the Control Group: After each game, we will reassign group members to the 3-person group. Namely, in each game you will be voting with different group members and you will not know who they are. We will not reveal the votes of anyone at any time during or after the experiment, so no one will know what other participants voted for.

Read to the Treatment Group: After each game, we will reassign group members to the 3-person group. Namely, in each game you will be voting with different group members and you will not know who they are. After all games are finished, we will post the "result sheet" at the entrance of the room for participants to view. The result sheet has all participants' votes in all games on it, which means your vote will be known to everyone at the end of the experiment.

Let us play a "Practice Game" now. It is "practice" in the sense that we will provide more explanations, but we will still make cash payments according to the voting outcome of your group. The 3 experimenters will now explain the game rules again individually (experimenters distribute Plates for Game 0, and explain rules again to OPs, Relatives, and Others respectively). *Explanation script*:

- Members of each 3-person group are randomly assigned; No one knows who the other members are;
- Indicate payoffs to "you" and "others" on both plates;
- Majority rule: If two other members voted for Blue while you voted for Red, you will get the allocation on the Blue Plate;
- Talking is prohibited. The plates you get are probably different from your neighbor's so there's no point talking or peeking;
- We will pay cash after each game;

We will now collect votes. Please do not say anything. Use your fingers to point out the plate of your choice (Red or Blue). The experimenters will write down each participant's vote and calculate outcome for each group using the majority rule.

Experimenters calculate results and pay out cash.

We have now finished the Practice Game. Please let us know if you have questions.

Now let us proceed to the 6 formal games of Round 1.

<u>Round 2</u>

We now proceed to **Round 2** of the experiment. Different from the group assignment rule of Round 1, villagers that brought in relatives or friends will be assigned to the same group as their relatives/friends throughout the experiment, while the third member of the 3-person group will be randomly assigned before each game. Namely, the structure of any 3-person group is:

You Your Relative/Friend A Third Person (Unknown)

Similar to Round 1, each participant will see a Red Plate and a Blue Plate in each game. You will still see yellow labels indicating ownership, but they are slightly different from Round 1. The row that is labelled "You" shows your payoffs in Red and Blue. The row labelled "Your Relative" shows payoff to your relative or friend in Red/ Blue, while the row labelled "Other" indicates payoff to the third unknown member in your group. We have also included the numerical values of payoffs in the brackets underneath the yellow labels.

We still follow the majority rule to determine voting outcome for each group. Namely, if two or more than two members in a 3-person group voted for Red, then for this group we will implement the allocation plan indicated by the Red plate. Similarly, if two or more members in a 3-person group voted for Blue, then for this group we will implement the Blue plate. After each game, the experimenters will calculate votes and distribute payoff in cash to each participant.

Read to the Control Group: Same as Round 1, we will not reveal the votes of anyone at any time during the experiment or after the experiment, so no one will know what other participants voted for.

Read to the Treatment Group: Same as Round 1, after all games are finished, we will post the "result sheet" at the entrance of the room for participants to view. The result sheet has all participants' votes in all games on it, which means that your vote will be known to your relative/friend and the other member at the end of the experiment.

We start Round 2 with a "Practice Game". Again, it is "practice" in the sense that we will provide more explanations, but we will still make cash payments according to the voting outcome of your group. The 3 experimenters will now explain the game rules individually (experimenters distribute Plates for Game 0, and explain rules again to OPs, Relatives, and Others respectively). *Explanation script*:

- Members of each 3-person group are: You, your relative/friend, and a third member that will be randomly assigned every game;
- Indicate payoffs to "you", "relative" and "other" on both plates;
- Majority rule: If two other members voted for Blue while you voted for Red, you will get the allocation on the Blue Plate;
- We will pay cash after each game;

We will now collect votes. Please do not say anything. Use your fingers to point out the plate of your choice (Red or Blue). The experimenters will write down each participant's vote and calculate outcome for each group using the majority rule.

Experimenters calculate results and pay out cash.

We have now finished the Practice Game. Please let us know if you have questions.

Now let us proceed to the 6 formal games of Round 2.

<u>Questionnaire</u>

Experimenters distribute questionnaires

A3. Summary Statistics By Village

Table C.1: Summary Statistics by Village

	Village 1	Village 2	Village 3	Village 4
Gender	0.70	0.83	0.56	0.59
Age	39.11	42.67	47.44	41.41
Education	8.08	7.72	7.56	8.00
Income	1939.33	816.67	1431.48	3055.56
Name Percentage	0.20	0.14	0.30	0.25
N	90	18	27	27

Part B: Supplementary Analysis

B1. Balance Check

Table C.2: Covariate Balance (Control vs Treatment)

	(1)		(2)		(3)	
	Control		Treatment		Difference	
	Mean	SD	Mean	SD	Difference	t-Statistics
Gender	0.64	0.48	0.71	0.46	-0.06	(-0.86)
Age	42.92	10.46	39.22	9.75	3.70*	(2.31)
Education	7.84	2.43	8.06	2.72	-0.22	(-0.53)
Income	1929.78	1924.64	1898.61	3728.06	31.16	(0.07)
Name Percentage	0.20	0.27	0.23	0.29	-0.03	(-0.64)
N	90		72		162	

B2. Game-by-Game Comparison

Figure C.1: Game 1 Variant 1

Game1 Variant 1			***Round 1***		***Round 2***	
			(1)	*(2)*	*(3)*	*(4)*
	Yes	No	No Pres	Pressure	No Pres	Pressure
Person 1	2	5	80.00%	87.50%	73.33%	83.33%
Person 2	8	5	63.33%	58.33%	66.67%	79.17%
Person 3	5	5	60.00%	66.67%	70.00%	58.33%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
7.50%	0.4719	10.00%	0.3893	-6.67%	0.4888	-4.17%	0.6643
-5.00%	0.7144	12.50%	0.3173	3.33%	0.7450	**20.83%**	**0.0961**
6.67%	0.6221	-11.67%	0.3819	10.00%	0.2638	-8.33%	0.5385

Figure C.2: Game 1 Variant 2

Game1 Variant 2			***Round 1***		***Round 2***	
			(1)	*(2)*	*(3)*	*(4)*
	Yes	No	No Pres	Pressure	No Pres	Pressure
Person 1	3	5	73.33%	79.17%	80.00%	91.67%
Person 2	7	5	53.33%	54.17%	66.67%	87.50%
Person 3	5	5	76.67%	66.67%	66.67%	79.17%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
5.83%	0.6262	11.67%	0.2384	6.67%	0.4888	12.50%	0.1853
0.83%	0.9525	**20.83%**	**0.0777**	13.33%	0.2550	**33.33%**	**0.0082**
-10.00%	0.4245	12.50%	0.3173	-10.00%	0.1841	12.50%	0.1841

Figure C.3: Game 1 Variant 3

Game1 Variant 3			***Round 1***		***Round 2***	
			(1)	(2)	(3)	(4)
	Yes	No	No Pres	Pressure	No Pres	Pressure
Person 1	4	5	70.00%	83.33%	73.33%	91.67%
Person 2	6	5	53.33%	66.67%	70.00%	79.17%
Person 3	5	5	76.67%	62.50%	66.67%	75.00%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
13.33%	0.2632	**18.33%**	**0.0878**	3.33%	0.7450	8.33%	0.4259
13.33%	0.3310	9.17%	0.4545	16.67%	0.2018	12.50%	0.2656
-14.17%	0.2657	8.33%	0.5143	-10.00%	0.3256	12.50%	0.2656

Figure C.4: Game 2 Variant 1

Game2 Variant 1			***Round 1***		***Round 2***	
			(1)	(2)	(3)	(4)
	Yes	No	No Pres	Pressure	No Pres	Pressure
Person 1	3	5	73.33%	66.67%	73.33%	70.83%
Person 2	5	5	46.67%	62.50%	73.33%	70.83%
Person 3	10	5	23.33%	29.17%	40.00%	50.00%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
-6.67%	0.6021	-2.50%	0.8422	0.00%	1.0000	4.17%	0.5748
15.83%	0.2545	-2.50%	0.8422	**26.67%**	**0.0029**	8.33%	0.4912
5.83%	0.6346	10.00%	0.4719	16.67%	0.1340	20.83%	0.1703

Figure C.5: Game 2 Variant 2

Game2 Variant 2				**Round 1**		**Round 2**	
				(1)	*(2)*	*(3)*	*(4)*
	Yes	No		No Pres	Pressure	No Pres	Pressure
Person 1	9	5		46.67%	29.17%	43.33%	41.67%
Person 2	5	5		46.67%	66.67%	36.67%	41.67%
Person 3	4	5		36.67%	66.67%	56.67%	75.00%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
-17.50%	0.1969	-1.67%	0.9043	-3.33%	0.7868	12.50%	0.3277
20.00%	0.1471	5.00%	0.7144	-10.00%	0.3746	**-25.00%**	**0.0558**
30.00%	**0.0285**	18.33%	0.1670	**20.00%**	**0.0314**	8.33%	0.5385

Figure C.6: Game 2 Variant 3

Game2 Variant 3				***Round 1***		***Round 2***	
				(1)	*(2)*	*(3)*	*(4)*
	Yes	No		No Pres	Pressure	No Pres	Pressure
Person 1	7	5		33.33%	20.83%	23.33%	33.33%
Person 2	5	5		23.33%	33.33%	30.00%	29.17%
Person 3	6	5		23.33%	16.67%	13.33%	25.00%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
-12.50%	0.3173	10.00%	0.4245	-10.00%	0.2638	**12.50%**	**0.0830**
10.00%	0.4245	-0.83%	0.9481	6.67%	0.6015	-4.17%	0.7466
-6.67%	0.5543	11.67%	0.2814	-10.00%	0.3256	8.33%	0.3277

Figure C.7: Game 2 Variant 4

Game2 Variant4				***Round 1***		***Round 2***	
				(1)	*(2)*	*(3)*	*(4)*
	Yes	No		No Pressure	Pressure	No Pressure	Pressure
Person 1	2	5		83.33%	66.67%	66.67%	83.33%
Person 2	8	5		66.67%	62.50%	76.67%	62.50%
Person 3	8	5		60.00%	50.00%	60.00%	33.33%

(2)-(1)		*(4)-(3)*		*(3)-(1)*		*(4)-(2)*	
Change	p-value	Change	p-value	Change	p-value	Change	p-value
-16.67%	0.1605	16.67%	0.1711	**-16.67%**	**0.0960**	16.67%	0.1617
-4.17%	0.7556	-14.17%	0.2657	10.00%	0.3746	0.00%	1.0000
-10.00%	0.4719	**-26.67%**	**0.0526**	0.00%	1.0000	-16.67%	0.1617

B3. Coefficient of Human Relationship

Relationship	Coefficient
Spouses	1
Parent and offspring	0.5
Parent in-law and offspring in-law	0.5
Grand-parent and grand-son/daughter	0.25
Siblings	0.25
In-law siblings	0.25
Parent's siblings/ offspring's siblings	0.125
Spouse's sibling's offspring	0.125
Cousins	0.0625
Non-relatives (neighbors, friends, etc.)	0