

7-2-2021

## Essays on Migration, Entrepreneurship, and Risk Hedging

Yiming Liu  
yliu074@fiu.edu

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>



Part of the [Economics Commons](#)

---

### Recommended Citation

Liu, Yiming, "Essays on Migration, Entrepreneurship, and Risk Hedging" (2021). *FIU Electronic Theses and Dissertations*. 4726.

<https://digitalcommons.fiu.edu/etd/4726>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

ESSAYS ON MIGRATION, ENTREPRENEURSHIP, AND RISK HEDGING

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

ECONOMICS

by

Yiming Liu

2021

To: Dean John F. Stack, Jr.  
Steven J. Green School of International and Public Affairs

This dissertation, written by Yiming Liu, and entitled Essays on Migration, Entrepreneurship, and Risk Hedging, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

---

Sheng Guo

---

Qiang Kang

---

Norihiko Matsuda

---

Cem Karayalcin, Major Professor

Date of Defense: July 2, 2021

The dissertation of Yiming Liu is approved.

---

Dean John F. Stack, Jr.  
Steven J. Green School of International and Public Affairs

---

Andrés G. Gil  
Vice President for Research and Economic Development  
and Dean of the University Graduate School

Florida International University, 2021

© Copyright 2021 by Yiming Liu

All rights reserved.

DEDICATION

To my parents.

## ACKNOWLEDGMENTS

I would like to take this opportunity to express my gratitude to my major Professor Cem Karayalcin as well as the members in my committee: Professor Sheng Guo, Professor Qiang Kang, and Professor Norihiko Matsuda. I have learned so much from them which can't be learned from textbooks. Without their help, I wouldn't be able to make it this far in my research. I would also like to thank Professor Cassian D'Cunha and Jonathan Casco from Instructional & Research Computing Center (IRCC) who assisted me along the way of data scraping using resources from HPC. I would like to thank Professor James Naismith, who invented basketball—one of the greatest sports in the world. Last but not the least, I would like to thank my parents who support me all the time.

ABSTRACT OF THE DISSERTATION  
ESSAYS ON MIGRATION, ENTREPRENEURSHIP, AND RISK HEDGING

by

Yiming Liu

Florida International University, 2021

Miami, Florida

Professor Cem Karayalcin, Major Professor

Expectations of higher income and better living conditions are the major causes behind migration. Whether migration is temporary or permanent is a consequence of both these expectations as well as the policies and the economic development of the host area. Migration comes with risks and migrants are more willing to take risks than nonmigrants. Entrepreneurship, like migration, also requires risk-taking for potentially high returns. The theme that connects the three chapters in the dissertation is the risks people take as they decide to migrate, to start new businesses, and the hedging strategies they adopt to mitigate these risks.

## INTRODUCTION

The transition from labor-intensive to capital-intensive industry in all sectors of China has been taking place under structural transformation. However, migrant workers are unprepared because of their low level of education. In the first chapter, the research focus is to figure out if return migrants are more influential than current migrants on the entrepreneurship of their nonmigrant family members. This chapter contributes to the literature on internal migration by studying the possibility of experience and resource sharing between migrants and nonmigrants. Nonmigrants have advantages over migrants in that they have more access to local resources while they are more credit-constrained and less experienced in the job market. Current migrants have more connections in the cities and earn more than return migrants. The empirical models utilized include a two-stage least square, a probit model with instrumental variable, and a recursive bivariate probit model. The endogeneity issue of return decision has been tackled by using the number of migrant workers who returned in 2013 out of the total number of migrant workers so far at the county level. The major conclusion is that current migrants are more influential than return migrants on the entrepreneurship of nonmigrants.

Entrepreneurs deal with risks from domestic markets, while small and medium-sized companies and large corporations that engage in international trade confront risks from both domestic and foreign markets. The second chapter utilizes data scraping to parse data from SEC EDGAR to study the effectiveness of the hedging strategies of trading companies in the U.S. markets in 2019. With both data scraping and textual analysis, we are able to quantify the keywords of the risk aversions and hedging strategies that appeared throughout the 10-k filing of each company. Market data and capital returns of each company have been used to get the expected exchange rate exposure, which is used to analyze the effectiveness of hedging strate-



gies of each company. This chapter contributes to the literature on risk management by combining both data scraping and textual analysis. The proportion of foreign sales out of total sales of each company is directly correlated with the exchange rate exposure of a company. Hedging strategies are effective in mitigating exchange rate exposure in the service sector.

The diversity of the language and religion in India renders a different environment for internal migration than it is in China. The Indian government has outlawed the caste system and the infrastructure in the major cities has been improved. The above-mentioned progress facilitates inter-state migration. We are interested in figuring out if inter-state migrants in India take advantage of this opportunity and are more active in the job market than their intra-state counterparts. This chapter contributes to the literature on internal migration in India by studying if there are more job opportunities available for inter-state migrants in other states. Three empirical models have been utilized in the analysis: a two-stage least square, a probit model with instrumental variable, and a recursive bivariate probit model. The instrumental variable used to tackle the endogeneity issue of the migration decision is the total number of households of inter-state migrants in 2009 over the total number of households of inter-state migrants at the district level. Inter-state migrants do a better job than their intra-state counterparts in the job market participation, especially in being day laborers. What's more, female inter-state migrants are more active than their intra-state and male counterparts in the job market participation.

## TABLE OF CONTENTS

CHAPTER	PAGE
1. Migration and Entrepreneurship in Rural China . . . . .	1
1.1 Introduction . . . . .	1
1.2 Literature Review . . . . .	6
1.3 Data and Descriptive Statistics . . . . .	8
1.4 Empirical Methods and Identification Strategy . . . . .	13
1.5 Regression Results . . . . .	16
1.6 Instrument Relevance and Exogeneity Checks . . . . .	20
1.7 Conclusion . . . . .	25
1.8 Appendix . . . . .	26
2. Risk Hedging of Trading Companies in the U.S. Markets . . . . .	36
2.1 Introduction . . . . .	36
2.2 Literature Review . . . . .	37
2.3 Descriptive Statistics . . . . .	38
2.4 The Theoretical Model . . . . .	43
2.5 The Empirical Model . . . . .	48
2.6 Regression Results . . . . .	50
2.7 Robustness Checks . . . . .	53
2.8 Conclusion . . . . .	55
2.9 Appendix . . . . .	56
3. Migration in India under Structural Transformation . . . . .	65
3.1 Introduction . . . . .	65
3.2 Literature Review . . . . .	66
3.3 Descriptive Statistics . . . . .	68
3.4 Empirical Models and Identification Strategy . . . . .	72
3.5 Regression Results . . . . .	75
3.6 Exogeneity Checks and the Relevance of Instrumental Variable . . . . .	78
3.7 Conclusion . . . . .	79
3.8 Appendix . . . . .	80
BIBLIOGRAPHY . . . . .	88
VITA . . . . .	92

## LIST OF TABLES

TABLE	PAGE
1.1 Migrants by working locations . . . . .	4
1.2 Distributions of age groups of migrants . . . . .	5
1.3 Education levels of migrant workers . . . . .	5
1.4 Industrial distributions of migrant workers . . . . .	5
1.5 Descriptive Statistics . . . . .	12
1.6 Return migrants on non-agricultural self-employment of nonmigrants . .	19
1.7 Instrument Relevance . . . . .	24
1.8 Signs and significance of regression results . . . . .	24
1.9 Return migrants on general labor force participation of nonmigrants . .	26
1.10 Return migrants on agricultural self-employment of nonmigrants . . . .	27
1.11 Return migrants on being homemakers of nonmigrants . . . . .	27
1.12 Return migrants on full-time study of nonmigrants . . . . .	28
1.13 Return migrants on general labor force participation of nonmigrants at or under 50 . . . . .	28
1.14 Return migrants on detailed labor force participation of nonmigrants at or under 50 . . . . .	29
1.15 Return migrants on general employment of nonmigrants at or under 50 .	29
1.16 Return migrants on detailed employment of nonmigrants at or under 50	30
1.17 Return migrants on non-agricultural self-employment of nonmigrants at or under 50 . . . . .	30
1.18 Return migrants on being homemakers of nonmigrants at or under 50 .	31
1.19 Return migrants on full-time study of nonmigrants at or under 30 . . . .	31
1.20 Return migrants on labor force participation of male nonmigrants . . . .	32
1.21 Return migrants on general employment of male nonmigrants . . . . .	32
1.22 Return migrants on non-agricultural self-employment of male nonmigrants	33
1.23 Return migrants on being homemakers of male nonmigrants . . . . .	33

1.24	Return migrants on full-time study of male nonmigrants . . . . .	34
1.25	Return migrants on non-agricultural self-employment of female nonmigrants . . . . .	34
1.26	Return migrants on being homemakers of female nonmigrants . . . . .	35
1.27	Return migrants on full-time study of female nonmigrants . . . . .	35
2.1	Descriptive Statistics . . . . .	43
2.2	Sectors . . . . .	56
2.3	Sector continues . . . . .	57
2.4	Distribution of hedging strategies by sector . . . . .	58
2.5	Distribution of total asset . . . . .	58
2.6	Regression results: all sectors, 3 years . . . . .	59
2.7	Regression results: manufacturing and service, 3 years . . . . .	59
2.8	Regression results: all sectors, 5 years . . . . .	60
2.9	Regression results: manufacturing and service, 5 years . . . . .	60
2.10	Regression results: all industry, 3 years controlling TA . . . . .	61
2.11	Regression results: manufacturing and service, 3 years controlling TA . . . . .	61
2.12	Regression results: all industries, 3 years controlling TA . . . . .	62
2.13	Regression results: manufacturing and service, 5 years controlling TA . . . . .	62
2.14	Regression results: total assets under 151 billion, 3 years . . . . .	63
2.15	Regression results: total assets under 151 billion, 5 years . . . . .	64
3.1	Migration status by age . . . . .	80
3.2	Migration status by education . . . . .	80
3.3	Migration status by employment . . . . .	81
3.4	Migration status by sectors . . . . .	81
3.5	Labor force participation . . . . .	82
3.6	Wage-employment . . . . .	82

3.7	Day laborer . . . . .	83
3.8	Unpaid family workers . . . . .	83
3.9	IV relevance . . . . .	84
3.10	Summary of Signs of Robustness Checks . . . . .	85
3.11	Female Labor Force Participation . . . . .	86
3.12	Female Wage-employment . . . . .	86
3.13	Female Day Laborer . . . . .	87

## CHAPTER 1

### Migration and Entrepreneurship in Rural China

#### 1.1 Introduction

As the structural transformation of both the developed and developing economies continues, capital and technology-intensive industries that require educated and skilled workers replace the traditional labor-intensive industries. On the one hand, emerging industries are developing at an unprecedented rate, whereas on the other, the average level of education among migrant workers in China falls behind this trend.<sup>1</sup>

A starting point of this paper is the unbalanced growth theory,<sup>2</sup> which states that different sectors expand at different rates over time. China invested heavily in manufacturing during the first few decades of its growth and yet initiated significant investment in agriculture only at the beginning of the 21<sup>st</sup> century. Xiwen Chen, an expert on agricultural issues in China, estimated that from 1953 to 1985, a total of 600 to 800 billion yuan<sup>3</sup> had been transferred from rural to urban areas. The Chinese government implemented the Household Registration System (hukou) in the late 1950s, restricting movement between urban and rural areas. Rural residents do not have access to the social welfare system including subsidized health care and free education in the city, and it is hard for them to purchase properties due to credit constraints and place of origin. After “Reform and Opening Up” at the end of the 1970s and with the gradually increasing demand for labor in the cities, rural

---

<sup>1</sup>According to the report from National Bureau of Statistics (NBS), from 2013 to 2018, there are over 55% of migrant workers whose highest level of education is middle school.

<sup>2</sup>Representative scholars of unbalanced growth theory are Albert O. Hirschman, Hans Singer, Paul Streeten, etc.

<sup>3</sup>GDP in China in 1953 and 1985 are 82.44 and 909.89 billion yuan respectively.

residents were allowed and started to work in the cities. However, they can only work in the cities without entitlement to such things as their children's education and home mortgages available to their urban counterparts (Chan, 2010). Migrant workers in China are those who work at a location other than their place of origin at the county level. With the number of migrant workers increasing, they have become the backbone of China's economic growth, and even of the entire world, especially in the construction and manufacturing<sup>4</sup> sectors.

Migrant workers who are from rural areas have made their contributions to the prosperity of the nation, and the Chinese government is now working on improving the living standards of rural residents. Migrant workers choose to work in the cities because of higher income and better living conditions. For migrant workers, the dilemma is that while settling down permanently in the city requires strenuous effort, there are sharp contrasts between urban and rural areas in terms of housing, education, and health care, which renders moving back a hard decision to make.

In general, migrant workers earn much less than their urban counterparts while earning much more than their rural counterparts, this being especially the case for wage workers. With more of the workforce in the cities, however, there remains unutilized potential for economic growth in rural areas. A more developed rural economy will attract more migrant workers back. With more return migrants and with improved infrastructure as well as investment incentives in the rural areas, more private investment will be made and more factories and companies will be established. Gradually, the gap between rural and urban areas will be eliminated.

This paper focuses on how return migrants and current migrants in the households affect the decision-making of their nonmigrant family members regarding labor

---

<sup>4</sup>Migrant workers are the major forces in those OEM (original equipment manufacturer) factories, as well as export-oriented factories.

force participation, especially self-employment. Return migrants are those who used to work in the cities and now live and work in their rural places of origin. Current migrants are those whose places of origin are rural areas and who now live and work in the urban areas.

According to the Report on Migrant Workers by the National Bureau of Statistics of China (NBS), the total number of migrant workers was 268.94 million in 2013, which had increased by 2.4% compared with the previous year, and among which 102.84 million worked in the vicinity of their home villages, with an increase rate of 3.6% compared with the previous year. In 2018, the total number of migrant workers was 288.36 million, which had increased by 0.6% compared with the previous year, while the rate of growth had decreased by 1.1%. Among the migrants, 115.7 million worked in the vicinity of their places of origin, with an increase rate of 0.6% compared with the previous year. The once-popular destination for the migrant workers—the city, was less popular in 2018, and there was a drop of 1.5% in the migrant population from 137.1 million to 135.06 million (Table 1.1).

To have a better understanding of migrant workers in China, one needs to look at their age groups, their level of education, and the sectors worked in 2013 and 2018. In 2013, 30.8% of the migrants were aged between 21 and 30, which was the largest age group, followed by those who were aged between 41 and 50, which made up 26.4% of the entire migration population. The middle-aged outnumbered other age groups. In 2018, the age group that made up the majority was those between 41 and 50, followed by those aged between 21 and 30, with percentages of 25.5% and 25.2% respectively. Compared with 2013, the total number of young migrant workers dropped while the middle-aged increased. (Table 1.2)

In 2013, 60.6% of migrant workers finished only middle school, and the illiteracy rate was 1.2%. The ones with associate and higher degrees accounted for 6.7%. In



2018, The illiteracy rate remained the same as in 2013, while those who finished only middle school made up the highest percentage — 55.8% — of the total migrant population, and those with associate and higher degrees were 10.9% of the population (Table 1.3).

In 2013, 56.8% of migrant workers worked in the secondary sector, among which 31.4% worked in the manufacturing sector, while 22.2% worked in the construction sector. 42.6% of migrant workers worked in the tertiary sector, among which 11.3% worked in the retail and wholesale sector, and 10.6% worked in the service sector. In 2018, the percentage of migrant workers who worked in the secondary sector was 49.1%, not as high as it was in 2013, and for the manufacturing and construction sectors, the figures dropped to 27.9% and 18.6% respectively. The number of migrant workers out of the entire migration population who worked in the retail and wholesale as well as the service sector increased to 12.1% and 12.2% respectively. (Table 1.4)

The year 2013 had seen a continuance of the decline of the traditional manufacturing sector in China, which resulted in temporary involuntary unemployment, with migrant workers being the major source of labor supply to the sector. While the total number of migrant workers increased by 1.7% (2.74 million) compared with the previous year, they had already faced the consequences of the structural transformation.

Table 1.1: Migrants by working locations

	2013		2018	
	Number (in million)	Annual Changes	Number (in million)	Annual Changes
Total Population	268.94	2.40%	288.36	0.60%
Work in Vicinity	102.84	3.60%	115.7	0.60%
Work in Cities	NA	NA	135.06	1.50%

Migrants are defined as people who work outside their places of residence.

Table 1.2: Distributions of age groups of migrants

	2013 (in %)	2018 (in %)
Age Groups		
16-20	4.7	2.4
21-30	30.8	25.2
31-40	22.9	24.5
41-50	26.4	25.5
50 above	15.2	22.4

Table 1.3: Education levels of migrant workers

	2013 (in %)	2018 (in %)
Highest Level of Education		
Illiterate	1.2	1.2
Primary School	15.4	15.4
Middle School	60.6	55.8
High School	16.1	16.6
Associate and above	6.7	10.9

Table 1.4: Industrial distributions of migrant workers

	2013 (in %)	2018 (in %)
<b>Primary Industry</b>	0.6	0.4
<b>Secondary Industry</b>	56.8	49.1
Manufacturing	31.4	27.9
Construction	22.2	18.6
<b>Tertiary Industry</b>	42.6	50.5
Retail and Wholesale	11.3	12.1
Transportation, Logistics, and Postal Services	6.3	6.6
Hospitality and Catering	5.9	6.7
Household Services	10.6	12.2

The rest of the chapter is structured as follows: section 1.2 is a brief summary of the current literature regarding immigration and migration, section 1.3 introduces sources of the data set in more detail, and some variables used in the regression. Section 1.4 introduces the empirical models utilized for analyzing the impact that return migrants and current migrants have on nonmigrants; section 1.5 shows the main results from the regressions. Section 1.6 summarizes regression results with different dependent variables, i.e., the impact from the return migrants on the decision of nonmigrants regarding their labor force participation, employment, unemployment, on self-employment. Section 1.7 concludes.

## **1.2 Literature Review**

This paper has been inspired by the work of Marchand, Naudé, and Siegel (2017), who studied the relationship between within-border and cross-border migration and entrepreneurship, and classified the relevant literature as follows: analysis of the entrepreneurship decision of immigrants and natives; remittances and entrepreneurship in home countries; return migrants and nonmigrant entrepreneurs. Here, we focus on the last of these issues. Because of differences in socio-economic, political, as well as cultural backgrounds, the factors that affect the entrepreneurship decision of migrants and nonmigrants tend to be different, though, there exists some common ground as well. According to Arif and Irfan (1997), and Piracha and Vadean (2010), who studied Pakistan and Albania, the number of self-employed return migrants outnumbered nonmigrants. What's more, based on the study from Black and Costaldo (2009), for those return migrants who had enough savings and lived abroad long enough, there was a high probability to start their own businesses. They also emphasized that this phenomenon was stronger among the poorer population.

Démurger and Xu (2010) found a similar trend among return migrants in China's Wuwei county: those return migrants with more accumulated savings and more frequent job changes were more likely than nonmigrants to become entrepreneurs. The case in China doesn't follow a universal pattern. For example, in some provinces like Zhejiang, where E-commerce is much more prevalent than other districts in China, local entrepreneurs performed better than return migrant entrepreneurs during the Great Recession, which is an interesting phenomenon that needs further analysis.

Djankov, Qian, Ronald, and Zhuravskaya (2006) carried out their research on China's entrepreneurs and answered the following question: what is the role that the economic and political environment, as well as legal institutions play in entrepreneurship? They have found that credit constraint, social networks, and individual characteristics are among the determinants of entrepreneurship. One thing worth mentioning is that academic achievements at school were uncorrelated with the entrepreneurship of migrants. Frjters, Kong, and Meng (2011) based their study on RUMiC (Rural-Urban Migrant Workers in China) in 2007 and 2008 and found that credit constraint was one of the hindrances for migrant workers to become entrepreneurs. What's more, they were more credit-constrained than their urban counterparts, and the overcoming of which may increase direct earnings which worth 2% of the GDP of China in 2008. They have also found that a 1% increase in the gap between the earnings of migrants and Urban Hukou holders in the salaried sector increased the proportion of wage-earning migrants who want to be entrepreneurs by 0.4%.

Chan (2010) also studied the migrant workers during the periods as Frjters et al. (2011) did, and worked on the interaction of China's economy with the world, especially through export. Chan drew the conclusion drawn that going back to their hometowns will make those migrants worse off.

In this paper, we try to study the behavior of migrant workers under temporary involuntary unemployment, partly due to structural transformation, and how their return to places of origin influences the self-employment decision of nonmigrant family members.

### **1.3 Data and Descriptive Statistics**

To track the dynamics of income distribution in China, five waves of household surveys had been conducted in 1989, 1996, 2003, 2008, and 2013, which was called the Chinese Household Income Project Survey (CHIPS). This project was initiated by a group of researchers at the Australian National University and Beijing Normal University, and was supported by the China National Bureau of Statistics (NBS) as well as the Institute for the Study of Labor (IZA). CHIPS 2013 covered statistics from 15 provinces in which there were 126 cities and 234 counties with a total number of 18,948 households. This study focused on the data of rural areas only. The rural data included a total number of 39,065 individual observations. For those household members who were too old or too young to take the survey or were away from homes, the questionnaire was answered by other family members. The questionnaire covered such basic information as age, gender, marital status, education, etc. What's more, there was a section about primary employment in 2013, which included such information as specific sectors of employment, income, fringe benefits, etc. For those who were self-employed, there were questions collecting information about the amount of investment needed to start their businesses, and the amount of loans borrowed if any, together with sources of funding. The section about migration experience provided readers with information about the migration experience of surveyees before 2013, what would be the plan for migration in the coming year, and

the reason(s) for choosing not to migrate. Moreover, there was a section about the financial status at the household level, which covered disposable income, deposits, debts, movable assets, and productive assets. In addition to the information about individuals and their households, there was some basic information about the parents of the household heads and their spouses, as well as the siblings of the surveyees. There was also information about land usage, demolition, and expropriation. The last two sections were about household borrowing and subjective questions about living standards respectively.

Descriptive statistics can be found in Table 1.5. The number in parenthesis is the total number of each subgroup. Based on migration experience, current location, as well as plan for migration in the following year, the sample was divided into three sub-samples, which were return migrant, current migrant, and nonmigrant respectively. The focus of this study is to compare the impact of return migrants with current migrants on the employment decision of their nonmigrant family members, especially their self-employment decisions. According to the employment status of 2013, each of the three sub-samples was further divided into being self-employed in agriculture<sup>5</sup> or being self-employed in the non-agricultural sector.<sup>6</sup> The average age of the current migrants who were non-agricultural entrepreneurs was about 5 years younger than their returned counterparts (36.13 vs. 41.28), while the average age of nonmigrants of the same type was the oldest, which was 44.8. The average age of nonmigrants who were self-employed in agriculture was the oldest, which was 50.05. Male current migrants who were self-employed in agriculture made up 82.59% of the sub-sample who were current migrants and who were self-employed in agriculture, which was the highest among all the male sub-samples, while their

---

<sup>5</sup>those who work on their own farmland

<sup>6</sup>One individual can work on more than one type of employment over the year.

nonmigrant counterparts made up 44.78% of the sub-sample who were nonmigrants and who were self-employed in agriculture, which was the lowest among all the male sub-samples. Current migrants who were self-employed in the non-agricultural sector had the longest years of schooling on average (8.46 years), followed by return migrants who were self-employed in the non-agricultural sector (8.16 years). Nonmigrants who were self-employed in agriculture received 6.41 years of education, which was the lowest. It takes 9 years for an individual in China to graduate from middle school. The top two groups whose highest level of education was middle school were current migrants who were self-employed in the non-agricultural sector, followed by return migrants who were self-employed in the non-agricultural sector (62.35% and 61.72%). 40.06% of nonmigrants who were self-employed in agriculture finished only middle school, which was the lowest. Current migrants who were self-employed in the non-agricultural sector ranked the top in high school or higher education, which was 16.75%. The sub-sample that ranked the bottom in high school or higher education was nonmigrants who were self-employed in agriculture, which was 7.10%. The following are highlights of the sectors with different entrepreneurs: 23.66% of return migrants who worked as non-agricultural entrepreneurs were in sales, which was the highest percentage level in this subgroup. The highest percentage level of non-agricultural entrepreneurs of current migrants was in sales as well (22.45%). The most popular sector among non-agricultural entrepreneurs of nonmigrants was also sales and there were 29.25% of nonmigrant entrepreneurs doing their businesses in this sector. The average proportion under household in the first column of Table 1.5 is the mean value of the number of household members that belong to each subgroup out of the total family members. Only the households that had family members that belonged to a specific subgroup were considered. On average, 57.03% of household members were self-employed in agriculture who had never migrated.

Self-employed return migrants and nonmigrants engaging in non-agricultural businesses on average made up 38.05% and 36.85% of the corresponding household respectively. Nonmigrants who were self-employed in agriculture ranked at the top of the average household income, which was 25,015.11 yuan, followed by 20,259.38 yuan, which was the average household income for non-agricultural entrepreneurs in the cities. Return migrants who start their own non-agricultural businesses reported an average household income of 16,442.91 yuan, which was more than the average household income of their counterparts who were self-employed in agriculture (13,441.24 yuan).<sup>7</sup> Nonmigrants who were self-employed in agriculture had the largest areas of land on average at the household level, and the smallest number went to return migrants who were self-employed in nonagricultural businesses.

---

<sup>7</sup>The average wage in China in 2013 was 52,388 yuan. Annual per capita disposable income in rural and urban areas were 9,429.6 and 26,467 yuan respectively. Source: National Bureau of Statistics of China



Table 1.5: Descriptive Statistics

Variables	Return Migrant (3221)		Current migrant (5982)		Non-migrant (25787)	
	Agri. self-emp (1781)	Non- agri. self- emp(499)	Agri. self- emp(988)	Non- agri. self- emp(448)	Agri. self- emp(11987)	Non- agri. self- emp(2379)
<b>Individual</b>						
Age	44.47	41.28	39.56	36.13	50.05	44.80
Male	69.68%	75.35%	82.59%	72.77%	44.78%	61.12%
Married	90.96%	91.18%	83.60%	83.26%	91.33%	93.61%
HH Head	56.99%	52.10%	55.16%	31.03%	40.07%	50.90%
schooling	7.32	8.16	7.88	8.46	6.41	7.94
Education Level	Total (1781)	Total (499)	Total (988)	Total (448)	Total (11977)	Total (2377)
No Sch	3.65%	1.20%	2.13%	2.46%	12.65%	3.45%
Primary Sch	31.78%	20.64%	21.46%	3.80%	39.10%	23.73%
Middle Sch	53.73%	61.72%	62.35%	59.38%	40.06%	55.41%
Tech. or Voc. Sch	1.35%	3.21%	1.41%	1.09%	1.09%	3.16%
High Sch or more	9.49%	13.23%	12.65%	16.75%	7.10%	14.26%
<b>Industry</b>	Total (1166)	Total (486)	Total (972)	Total (441)	Total (3502)	Total (2301)
Agriculture	10.98%	7.82%	2.78%	1.81%	17.79%	10.73%
Mining	2.74%	1.44%	3.60%	0.91%	2.43%	0.74%
Manufacturing	14.41%	7.20%	18.21%	9.30%	18.45%	12.43%
Energy	1.72%	1.03%	0.93%	0.91%	1.03%	0.83%
Construction	41.17%	19.75%	46.19%	18.37%	22.70%	8.82%
Sales	6.17%	23.66%	3.70%	22.45%	9.31%	29.25%
Transportation	5.75%	15.43%	5.66%	11.34%	5.17%	11.52%
Hospitality	3.17%	5.76%	4.22%	9.30%	4.05%	5.95%
IT	0.09%	0.21%	0.41%	0.91%	0.14%	0.17%
Finance	0.26%	0.21%	0.21%	0.45%	0.14%	0.09%
Real Estate	1.11%	0.62%	0.93%	0.23%	0.26%	0.26%
Leasing	0.43%	1.65%	0.72%	3.85%	0.63%	2.61%
Research	0.09%	0.00%	0.00%	0.00%	0.06%	0.04%
Infrastructure	0.51%	0.21%	0.41%	0.00%	0.83%	0.26%
Services	8.66%	11.32%	10.60%	17.46%	9.42%	12.34%
Education	0.34%	0.21%	0.31%	0.45%	1.03%	0.52%
Public Health	0.51%	0.41%	0.41%	1.13%	1.17%	1.87%
Entertainment	0.34%	1.03%	0.21%	1.13%	0.43%	0.70%
Social Welfare	1.54%	1.65%	0.51%	0.00%	4.97%	0.87%
Int'l Org.	0.00%	0.41%	0.00%	0.00%	0.00%	0.00%
<b>Household</b>						
HH included	1421	393	812	327	6877	1655
Avg. Proportion	0.393	0.3805	0.3548	0.3685	0.5703	0.4753
Income	13441.24	16442.91	13916.26	20259.38	19194.42	25015.11
Inc. per person	10779.69	12950.09	11469.36	15243.65	11031.24	17486.38
Land (mu)	2.06	1.30	2.08	1.64	3.75	1.63
Land per person	2.14	1.51	1.80	1.45	2.26	1.50

Source: CHIPS 2013. Return migrants are those who had migrated before 2013, and had no plan to migrate. Out-migrants are those who planned to migrate in few months, or those who had already migrated. Non-migrants are those who had never migrated.

Values are either in percentages or the average of the group for each column.

Avg. Proportion is the mean value of the proportion of family members that fit into each column.

## 1.4 Empirical Methods and Identification Strategy

Both return migrants and current migrants can influence the employment decisions of their nonmigrant family members. Return migrants have accumulated a certain amount of capital as well as working experience. Current migrants are similar to their returned counterparts except that they have access to resources and infrastructure in the cities that are not available to return migrants. After moving back, return migrants lose some of the connections they made in the cities.<sup>8</sup> Current migrants, especially those who plan to start their own businesses, have advantages over their returned counterparts in that they have easier access to tangible and intangible resources necessary to start and to run a business. Such essential resources include internet access, transportation, and policies that promote entrepreneurship, etc.<sup>9</sup>

To compare the influence of the return migrants with the current migrants on the entrepreneurship of their nonmigrant family members, this paper utilized three models: a two-stage least square (2SLS), a probit model with instrumental variable (IV probit), and a recursive bivariate probit (Recursive) model. We start with the observation that the return migration decision is not totally exogenous to the entrepreneurship decision. The decision to return is made under a variety of unobserved circumstances. What is correlated with the decision to move back and can be observed from the data is the number of migrant workers who returned in 2013 out of the total number of migrant workers so far at the county level,<sup>10</sup> which is

---

<sup>8</sup>Wahba and Zeno (2012) stated that there is a tradeoff between the accumulation of human and social capital and loss of social capital for return migrants.

<sup>9</sup>Chaurey and Le (2019) have found that improved and renovated infrastructure has positive influences on the set up of small businesses.

<sup>10</sup>Hereinafter referred to as the “the return ratio”.

used as the instrumental variable. The higher the return ratio in a county, the more likely the migrant workers will return to that county. The reasoning behind this is that when migrant workers move out to work, they get information about jobs from the townsfolk who are already in the cities. Advice from townsfolk is more likely to be accepted than information from other sources, especially when the number of townsfolk that provide information is large, a phenomenon known as the bandwagon effect. The same occurs with migrant workers in the cities whose decisions are influenced by people around them, especially those who come from the same village or county as they do. When migrant workers learn about the return migration of their townsfolk, especially in increasing numbers, more current migrants will follow. The identification strategy is used on 2SLS, IV probit, and Recursive model.

For the 2SLS we have

$$Y_i^* = a_0 + a_1 R_i^* + a_2 X_i + a_3 C_i + e_i \quad (1.1)$$

$$\hat{R}_i = \alpha_0 + \alpha_1 Rtn_i + \alpha_2 X_i + \alpha_3 C_i + \mu_i \quad (1.2)$$

$$Y_i^* = \beta_0 + \beta_1 \hat{R}_i + \beta_2 X_i + \beta_3 C_i + \epsilon_i \quad (1.3)$$

where

$Rtn_i$  denotes the return ratio.

$X_i$  denotes personal characteristics.

$C_i$  denotes the county level statistics.

$Y_i^*$  and  $R_i^*$  are unobserved latent variables with

$$Y_i = \begin{cases} 1 & \text{if } Y_i^* > 0 \\ 0 & \text{if } Y_i^* \leq 0 \end{cases} \quad (1.4)$$

and

$$R_i = \begin{cases} 1 & \text{if } R_i^* > 0 \\ 0 & \text{if } R_i^* \leq 0 \end{cases} \quad (1.5)$$

$Y_i^*$  and  $R_i^*$  are expected values which are obtained from the answers of the questionnaire: if an individual is a return migrant, and doesn't plan to migrate, it is regarded that the expected return from being a return migrant outruns the expected return of being a migrant worker. A dummy variable has been created from the questionnaire, return migrants equals 1, and 0 otherwise. Similar logic goes with the expected return to become an entrepreneur.  $Rtn_i$  denotes the instrumental variable to tackle the endogeneity issue.  $X_i$  denotes personal characteristics, and  $C_i$  denotes the county level statistics.  $ei$ ,  $\mu_i$ , and  $\epsilon_i$  are the error terms.

The IV probit model that takes into account the endogeneity issue analyzes the same variables as in the 2SLS model.

The recursive bivariate probit model is as follows:

$$Prob[Y = 1, R = 1 | \mathbf{X}_1, \mathbf{X}_2] = \phi(\theta_1 \mathbf{X}'_1 + \lambda R, \theta_2 \mathbf{X}'_2, \delta) \quad (1.6)$$

where  $\mathbf{X}_1$  and  $\mathbf{X}_2$  are the individual characteristics that influence the return and entrepreneurial decisions of migrant workers. According to Greens (1983), "the endogenous nature of one of the variables on the right-hand side of the first equation can be ignored in formulating the log-likelihood." This model is recursive and simultaneous. After figuring out how this decision-making works, the next step analyzes the likelihood for their nonmigrant family members to become entrepreneurs. The Econometric specification is as follows:

$$Y_i^* = \gamma_0 + \gamma_1 R_i + \gamma_2 X_i + \gamma_3 C_i + \delta_i \quad (1.7)$$

$$R_i^* = \rho_0 + \rho_1 Rtn_i + \rho_2 X_i + \rho_3 C_i + \tau_i \quad (1.8)$$

One of the assumptions of the recursive bivariate probit model is that the decisions on entrepreneurship and return are interdependent, with  $cov(\delta_i, \tau_i) = \delta$ . To obtain a more robust result to distributional misspecification, according to Monfardini and Radice (2008), an instrument  $Rtn_i$  is used, which is the same IV used in the previous models.

## 1.5 Regression Results

The empirical analysis starts by comparing the influence of return migrants with that of current migrants on the labor force participation of their nonmigrant household members, followed by comparing the influences of return migrants with current migrants on other employment types of their nonmigrant household members. The employment types are: being employed, being unemployed, being a homemaker<sup>11</sup>, being a family worker,<sup>12</sup> being a full-time student, and being self-employed in either agricultural or non-agricultural sectors. Return migrants are defined at the household level. If a household has one or more family member(s) who is/are return migrant(s), the variable *rmhh* equals 1, and vice versa. The dependent variable—self, is the entrepreneurial decision at the individual level. More specifically, the empirical analysis tries to tease out the impact of return migrants at the household level on the entrepreneurial decision of nonmigrant individuals in the household.<sup>13</sup> Households with both return migrants and current migrants are excluded from the

---

<sup>11</sup>Doing family chores of his/her own family.

<sup>12</sup>Someone who offers labor to the neighborhood during sowing, harvest, etc, or who helps with family chores of other families.

<sup>13</sup>Both household level and individual dummies are created based on the whole sample, and when running the regression, the sample is restricted to households with nonmigrants.

regression.<sup>14</sup> All three models (2SIS, IV probit, and Recursive model) have been utilized to analyze each employment type mentioned above.

Table 1.6 is the regression result with being self-employed in the non-agricultural sector as the dependent variable and being a return migrant at the household level as the independent variable. Return migrants are found to be less influential than current migrants when comparing the influence on the entrepreneurial decision of their nonmigrant family members in the non-agricultural sector. The regression results from the 2SLS, IV probit, and Recursive model are significant at 99% confidence intervals. A detailed analysis of the regression result can be found in the next paragraph.

Both return migrants and current migrants have accumulated working experience which is hard to get for nonmigrants. Moreover, both groups earn a higher salary than nonmigrant family members. Once the migrant workers return, however, their advantages of having access to modernized infrastructure<sup>15</sup> as well as social networks<sup>16</sup> are diminished. Having no or insufficient access to paved roads, electricity, and the internet makes it hard for return migrants to set up businesses. The likelihood for individuals to start their own businesses in general declines after a certain age.<sup>17</sup> The following is an analysis of the coefficients of the age and age

---

<sup>14</sup>Including or excluding households with both return migrants and current migrants barely change the result, and there are 802 households with both return migrants and out-migrants.

<sup>15</sup>Chaurey and Le (2019) have found that improving infrastructure is an effective way of stimulating local entrepreneurship.

<sup>16</sup>Wahba and Zenou (2012) proved in their paper that physical and social capital are complements during the temporary migration, once returned, the social networks accumulated would be lost.

<sup>17</sup>Lévesque and Minniti (2006) showed in their model that entrepreneurship started to decline after a certain age.

squared from the recursive bivariate probit model:  $0.117057 - 2*(0.00132)*age > 0$ ,<sup>18</sup> therefore, the value of age should be less than 44.34 for the equation to hold. The results from the 2SLS and IV probit model are 42.76 and 44.38 respectively. Migrants have accumulated the experience and wealth necessary to start their own businesses during and before middle age. What's more, individuals who are middle-aged or younger are more willing to take risks,<sup>19</sup> an essential characteristic of entrepreneurs. Current migrant entrepreneurs are on average four years younger than their return counterparts. They have access to social networks as well as the infrastructure necessary to start and to run a business. Being young and having access to resources necessary to become entrepreneurs, current migrants have more influence on the entrepreneurship of nonmigrant family members than return migrants.

---

<sup>18</sup>Entrepreneurship is positively correlated with age, as has been shown in the regression result.

<sup>19</sup>Djankov, Qian, etc., (2006) mentioned in one of their specifications that being risk-loving and greedy are consequential determinants of entrepreneurship in China.

Table 1.6: Return migrants on non-agricultural self-employment of nonmigrants

Non-agricultural self-employment									
Variables	2SLS			IV probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
<b>First Stage*</b>									
rtnpcthh	0.924678	***	0.048051	2.681104	***	0.153821	2.674624	***	0.153925
age	-0.00449	***	0.000764	-0.01338	***	0.002227	-0.01343	***	0.002241
agesqrd	3.95E-05	***	9.86E-06	0.000116	***	2.94E-05	0.000117	***	2.96E-05
perGDP	1.08E-07		7.10E-08	2.60E-07		2.4E-07	2.74E-07		2.37E-07
fisinc	1.78E-07	***	5.19E-08	5.30E-07	***	1.72E-07	5.12E-07	***	1.71E-07
fisspd	-1.45E-07	***	4.82E-08	-4.39E-07	**	1.61E-07	-4.22E-07	***	1.60E-07
ttlsale	-3.61E-10	***	2.03E-09	-1.86E-10	***	6.92E-09	-3.69E-10	***	6.94E-09
Cons	0.134721		0.025869	-1.0220	***	0.084162	-1.02232	***	0.084385
<b>Second Stage*</b>									
rmhh	-0.06712	***	0.026093	-0.57396	***	0.205973	-0.53178	***	0.176465
age	0.006798	***	0.000511	0.11717	***	0.009934	0.117057	***	0.009883
agesqrd	-8E-05	***	6.56E-06	-0.00132	***	0.000121	-0.00132	***	0.000121
perGDP	3.19E-08		3.50E-08	1.50E-07		2E-07	1.43E-07		2.04E-07
fisinc	8.66E-08	***	2.70E-08	6.36E-07	***	1.88E-07	6.37E-07	***	1.89E-07
fisspd	-5.56E-08	***	2.23E-08	-3.99E-07	**	1.63E-07	-4.00E-07	**	1.64E-07
ttlsale	-2.47E-09	***	8.36E-10	-2.66E-08	***	8.36E-09	-2.64E-08	***	8.27E-09
Cons	-0.01789		0.01158	-3.4909	***	0.247012	-3.50487	***	0.235587
Size	10114			10114			10114		
F(7, 186)	43.77								
Prob>F	0								
R-squared	0.0165								
Root MSE	0.23942								
Wald Chi2(7)				247.45					
Log pseudo				-8165.38			-7895.86		
rho							0.3735		0.1097
<b>Wald test</b>									
Chi2(1)				8.84			9.4854		
Prob>Chi2				0.0029			0.0021		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

The 1st stage for 2SLS and IV probit model is the 2nd stage for Recursive model, and the same for the 2nd. perGDP, fisinc, fisspd, ttlsale are per capita GDP, fiscal income, fiscal spending, and total sales.

The unit of perGDP is yuan, and the unit for the rest is at 10,000 yuan.



## 1.6 Instrument Relevance and Exogeneity Checks

The first stage result from the 2SLS and IV probit model and the second stage result from the Recursive model with the return decision as the dependent variable and the return ratio as the independent variable are shown in Table 1.7. Return decision is positively correlated with Return ratio from all three models at 99% confidence intervals.

The labor force participation of surveyees is based on their employment at the end of 2013. Those who were employed or temporarily unemployed at the end of 2013 were in the labor force, vice versa. Detailed labor force participation took into account all the employment information of each participant throughout the year. There were questions asking whether the participant had engaged in any agricultural work, whether the participants had engaged in any non-agricultural self-employment, whether the participant had engaged in any wage-employment, etc. The same measures have been taken when creating variables of employment and unemployment. Variables of self-employment, wage-employment, and other detailed employment types are based on the information of detailed labor force participation.

Throughout the rest of the section, the dependent variable of each regression is the return migration decision at the household level. Each regression is the comparison between the influence of return migrants and current migrants on the job search of nonmigrant family members. The regression starts from analyzing the whole sample, followed by analyzing surveyees who are 50<sup>20</sup> or younger, and male and female surveyees.

The signs and significance levels of all the regressions are shown in Table 1.8, and the rest of the regression results are in the appendix. Return migrants are more

---

<sup>20</sup>The threshold picked for the sub-sample of full-time study is at or under 30 years old.

influential on nonmigrant family members than current migrants in the following instances: general labor force participation, general labor force participation of nonmigrants at or below 50, detailed labor force participation of nonmigrants at or below 50, general employment of nonmigrants at or below 50, detailed employment of nonmigrants at or below 50, labor force participation of male nonmigrants, and general employment of male migrants. Current migrants have a stronger impact on nonmigrant family members than return migrants in the following cases: self-employment in agriculture, being homemakers and being full-time student, self-employment of nonmigrants at or below 50 in the non-agricultural sector, being homemakers of nonmigrants at or below 50, being full-time students of nonmigrants at or below 50, non-agricultural self-employment of male nonmigrants, being homemakers of male nonmigrants, being full-time students of male migrants, being self-employment in non-agricultural sector of female nonmigrants, being homemakers of female nonmigrants, being full-time students of female nonmigrants.

Tables 1.9 to 1.12 show the regression results when the dependent variable is general labor force participation, nonmigrant being self-employed in agriculture, nonmigrant working as a homemaker, and nonmigrant working as a full-time student respectively. Table 1.9 indicates that the influence of return migrants on their nonmigrant family members in the general labor force participation is insignificant from the 2SLS model, while it is significant at 90% and 95% confidence intervals from IV probit and Recursive model respectively. Table 1.10 shows that current migrants are more influential on the self-employment decision of nonmigrant family members in agriculture, however, results from other sub-samples are insignificant. Current migrants are more influential than return migrants on the decision-making of nonmigrants in becoming homemakers and full-time students as shown in Table 1.11 and 1.12 respectively.

Tables 1.13 to 1.19 show the regression results from the sub-sample of surveyees who aged at or below 50.<sup>21</sup> The regression result in Table 1.13 indicates that the influence of the return migrants on the general labor force participation of nonmigrant family members is positive, and the result is statistically significant at 99% confidence intervals from all three models. The result in Table 1.14 shows that return migrants have a significant positive impact on the detailed labor force participation of nonmigrant family members. Regression results in Table 1.15 indicate that households with return migrants have a larger impact on the general employment of nonmigrant family members. Return migrants are more influential on nonmigrant family members in terms of detailed employment, and the regression result (Table 1.16) is significant at 90%, 90%, and 95% confidence intervals from the 2SLS, IV probit, and Recursive models. The regression results in Table 1.17 show that current migrants are more influential than return migrants in the self-employment decision of nonmigrant family members in the non-agricultural sector. Current migrants are more influential than return migrants on nonmigrant family members in becoming homemakers and full-time students as shown in Table 1.18 and 1.19 respectively.

Tables 1.20 to 1.24 show the regression results from the male sub-sample. The result from the male sub-samples in Table 1.20 indicates that return migrants are more influential on the labor force participation of male nonmigrants. Return migrants also have a larger influence on male nonmigrants in terms of general employment (Table 1.21). Current migrants are more influential than return migrants on male nonmigrant family members in terms of self-employment in the non-agricultural sector (Table 1.22). Current migrants are also more influential on male nonmigrants in being homemakers and being full-time students (Tables 1.23 and 1.24 respectively).

---

<sup>21</sup>People aged at or below 50 are the backbones of the labor force, and their experience and expertise are more influential in the labor force.

Tables 1.25 to 1.27 show the regression results from the female sub-sample. Current migrants are more influential on female nonmigrants in self-employment in the non-agricultural sector (Table 1.25). Tables 1.26 to 1.27 show that current migrants have a more prominent impact than return migrants on female nonmigrants in being homemakers and full-time students.

Current migrants are not only more influential on nonmigrant family members in self-employment, but also in being homemakers and full-time students. Nonmigrants are more likely to become homemakers with current migrant family members. Families with current migrants report a higher income on average than those families with return migrants, therefore, their nonmigrant family members can postpone their job search. Current migrants are still working in the cities, and having access to first-hand information about school information as well as living conditions. What's more, they can offer emotional as well as financial support while those family members study in the cities.<sup>22</sup>

---

<sup>22</sup>Migration status are related to working experiences only, if someone who has never migrate out to work, their status of migration is nonmigrant.

Table 1.7: Instrument Relevance

Variables	rmhh								
	2SLS			IV probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
rtnpct	0.924678	***	0.048051	2.681104	***	0.153821	2.674624	***	0.153925
age	-0.00449	***	0.000764	-0.01338	***	0.002227	-0.01343	***	0.002241
agesqrd	3.95E-05	***	9.86E-06	0.000116	***	2.94E-05	0.000117	***	2.96E-05
perGDP	1.08E-07		7.10E-08	2.60E-07		2.40E-07	2.74E-07		2.37E-07
fisinc	1.78E-07	***	5.19E-08	5.30E-07	***	1.72E-07	5.12E-07	***	1.71E-07
fisspd	-1.45E-07	***	4.82E-08	-4.39E-07	***	1.61E-07	-4.22E-07	***	1.60E-07
ttlsale	-3.61E-10		2.03E-09	-1.86E-10		6.92E-09	-3.69E-10		6.94E-09
Cons	1.35E-01	***	0.025869	-1.02195	***	0.084162	-1.02232	***	0.084385
Sample	10151			10151			10114		
F(7, 186)	74.58								
Prob<math>\chi^2</math>F	0								
R-squared	0.0906								
Root MSE	0.44122								
Wald Chi2(7)				0					
Prob>chi2				0.0733			-7895.86		
Pseudo R2									
rho							0.3735		0.1097
<b>Wald test</b>									
Chi2(1)							9.4854		
Prob>Chi2							0.0021		

Table 1.8: Signs and significance of regression results

	lfp1	lfp2	emp1	emp2	unemp1	unemp2	selfag	selfnag	sch	empwg	hm	hw
<b>All</b>	+	+	+	+	+	-	-	-	-	+	-	-
2SLS							***	***	**		*	
IV probit	*						***	***	***		*	
Recursive	**						***	***	***		*	
<b>&lt;=50*</b>	+	+	+	+	+	-	+	-	-	+	-	+
2SLS	***	***	***	*				***	***		*	
IV probit	***	***	***	*				***	***		**	
Recursive	***	***	***	**				***	***		**	
<b>Male</b>	+	+	+	+	-	-	+	-	-	+	-	+
2SLS	*		*					*	*		*	
IV probit	***		***	**				**	***	*	*	
Recursive	**		***	***				**	***	**	**	
<b>Female</b>	+	+	+	+	+	-	-	-	-	+	-	-
2SLS								**	**		*	
IV probit								***	***		*	
Recursive								***	***		*	

\*: significant at 10%, \*\*: significant at 5%, \*\*\*: significant at 1%.

plus sign means nonmigrants are more likely to engage in each employment with return migrants at home. lfp1, emp1, and unemp1 are labor force participation, employment, and unemployment at the end of 2013. lfp2, emp2, and unemp2 take into account employment information throughout the year. selfag means self-employed in agriculture, and selfnag means self-employed in other industries. sch represents full-time study, empwg means wage-employed, hm is homemaker, and hw is homemaker. for sch, age is restricted to at or under 30 when studying sub-samples.

## 1.7 Conclusion

Having nonmigrant entrepreneurs or return migrant entrepreneurs at their place of origin doesn't significantly reduce the increasing income gap between rural and urban areas, but it is necessary to create an environment for further regional growth. The jobs created by both return entrepreneurs and nonmigrant entrepreneurs are relatively small in number compared to the supply of labor in the rural areas.<sup>23</sup> Because of credit constraints, the shortage of capital, and inadequate local infrastructure, it is difficult for return migrants to set up large-scale businesses.<sup>24</sup> The rural economy can catch up with the urban economy only when medium and large corporations establish themselves in rural areas. For this to happen, electrification, road connectivity, as well as internet access are prerequisites.

The government has been working on improving the living standards of the residents in rural areas. Investments are being made in rural areas to pave roads, to upgrade infrastructure, and to build new schools. However, the current gap between rural and urban areas is such that much more needs to be done.

---

<sup>23</sup>According to CHIPS 2013, rural entrepreneurs on average created 0.78 jobs which don't include hiring their own family members.

<sup>24</sup>From the data set, out of 39,065 observations, there are 2,037 who reported positive (others are either negative, 0, or NA.) total investment before the start of their businesses. The minimum is 8 yuan, the median is 20,000 yuan, the mean is 56,987.89 yuan, and the maximum is 2.1 million yuan.

## 1.8 Appendix

Table 1.9: Return migrants on general labor force participation of nonmigrants  
Labor force participation

Variables	2SLS		IV Probit		Recursive	
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.
rmhh	0.118893	0.074939	0.46147 *	0.256782	0.469973 **	0.232204
age	0.04438 ***	0.00122	0.192917 ***	0.007951	0.192748 ***	0.007975
agesqrd	-0.00048 ***	1.47E-05	-0.00204 ***	9.19E-05	-0.00204 ***	9.22E-05
perGDP	1.40E-07	1.11E-07	4.91E-07	3.64E-07	4.91E-07	3.63E-07
fisinc	-3.60E-08	7.22E-08	-1.83E-07	2.52E-07	-1.86E-07	2.51E-07
fisspd	2.29E-08	6.93E-08	1.29E-07	2.43E-07	1.32E-07	2.43E-07
ttlsale	-1.75E-09	2.49E-09	-5.33E-09	8.18E-09	-5.35E-09	8.18E-09
Cons	-3.64E-01 ***	0.036386	-4.00908 ***	0.13246	-4.00761 ***	0.1313
Sample	10056		10056		10056	
F(7, 186)	265.26					
Prob>F	0					
R-squared	0.3102					
Root MSE	0.41401					
Wald Chi2(7)			979.91			
Log pseudo			-10768.8		-10500.2	
rho					-0.3154	0.1359
<b>Wald test</b>						
Chi2(1)			3.92		4.6857	
Prob>Chi2			0.0476		0.0304	

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.10: Return migrants on agricultural self-employment of nonmigrants

Agricultural self-employed									
2SLS			IV Probit				Recursive		
Variables	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
rmhh	-0.06712	***	0.026093	-0.57396	***	0.205973	-0.53178	***	0.176465
age	0.006798	***	0.000511	0.11717	***	0.009934	0.117057	***	0.009883
agesqrd	-8E-05	***	6.56E-06	-0.00132	***	0.000121	-0.00132	***	0.000121
perGDP	3.19E-08		3.50E-08	1.50E-07		2E-07	1.43E-07		2.04E-07
fisinc	8.66E-08	***	2.70E-08	6.36E-07	***	1.88E-07	6.37E-07	***	1.89E-07
fisspd	-5.56E-08	***	2.23E-08	-3.99E-07	**	1.63E-07	-4.00E-07	**	1.64E-07
ttlsale	-2.47E-09	***	8.36E-10	-2.66E-08	***	8.36E-09	-2.64E-08	***	8.27E-09
Cons	-0.01789		0.01158	-3.4909	***	0.247012	-3.50487	***	0.235587
Size	10114			10114			10114		
F(7, 186)	43.77								
Prob>F	0								
R-squared	0.0165								
Root MSE	0.23942								
Wald Chi2(7)				247.45					
Log pseudo				-8165.38			-7895.86		
rho							0.3735		0.1097
<b>Wald test</b>									
Chi2(1)				8.84			9.4854		
Prob>Chi2				0.0029			0.0021		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.11: Return migrants on being homemakers of nonmigrants

Homemaker									
2SLS			IV Probit				Recursive		
Variables	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
rmhh	-0.11262	*	0.064645	-0.47582	*	0.277876	-0.47831	*	0.254848
age	0.009858	***	0.001091	0.080033	***	0.005713	0.079752	***	0.005718
agesqrd	-8.00E-05	***	1.31E-05	-0.00071	***	6.11E-05	-0.0007	***	6.08E-05
perGDP	-1.83E-07	**	9.05E-08	-9.24E-07	*	5.46E-07	-9.23E-07	*	5.41E-07
fisinc	6.49E-08		6.19E-08	2.88E-07		2.57E-07	2.91E-07		2.57E-07
fisspd	-5.46E-08		5.66E-08	-2.41E-07		2.36E-07	-2.44E-07		2.36E-07
ttlsale	3.70E-09	*	2.06E-09	1.44E-08	*	7.81E-09	1.43E-08	*	7.86E-09
Cons	-0.01679		0.028574	-2.69292	***	0.214067	-2.68345	***	0.209695
Sample	10114			10114			10114		
F(7,186)	44.07								
Prob>F	0								
R-squared	0.0513								
Root MSE	0.361								
Wald Chi2(7)				473.83					
Log pseudo				-10089.2			-9819.28		
rho							0.3084		0.1523
<b>Wald test</b>									
Chi2(1)				3.24			3.5882		
Prob>Chi2				0.0719			0.0582		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.



Table 1.12: Return migrants on full-time study of nonmigrants

Full-time Study									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.08588	**	0.04102	-0.86449	***	0.241788	-0.69244	***	0.195569
age	-0.02271	***	0.00098	0.66755	***	0.044563	0.688153	***	0.036585
agesqrd	0.000144	***	1.00E-05	-0.02528	***	0.001694	-0.02604	***	0.001402
perGDP	-7.79E-08		7.10E-08	-2.00E-07		2.87E-07	-2.09E-07		3.09E-07
fisinc	-9.28E-08	***	3.72E-08	-3.53E-07		2.75E-07	-4.20E-07	**	2.82E-07
fisspd	8.93E-08	***	3.28E-08	4.44E-07	*	2.59E-07	5.07E-07		2.65E-07
ttlsale	-2.15E-12		1.53E-09	-1.36E-09		7.19E-09	-1.39E-09		7.13E-09
Cons	8.30E-01	***	0.033821	-2.44958	***	0.283565	-2.60993	***	0.229069
Sample	10056			10056			10056		
F(7,186)	274.37								
Prob>F	0								
R-squared	0.4156								
Root MSE	0.31225								
Wald Chi2(7)				985.49					
Log pseudo				-7143.84			-6877.17		
rho							0.4472		0.11
<b>Wald test</b>									
Chi2(1)				11.93			12.2413		
Prob>Chi2				0.0006			0.0005		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.13: Return migrants on general labor force participation of nonmigrants at or under 50

Labor force participation at or under 50									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	0.177171	***	0.065084	0.73344	***	0.249094	0.705941	***	0.220628
age	0.024524	***	0.001978	0.279095	***	0.020502	0.279966	***	0.019926
agesqrd	-8.9E-05	**	3.75E-05	-0.00333	***	0.000294	-0.00334	***	0.000286
perGDP	9.61E-08		7.71E-08	3.17E-07		2.90E-07	3.17E-07		2.93E-07
fisinc	1.07E-07		7.36E-08	3.08E-07		2.94E-07	3.13E-07		2.92E-07
fisspd	-6.37E-08		6.73E-08	-1.28E-07		2.65E-07	-1.32E-07		2.64E-07
ttlsale	-5.64E-10		2.88E-09	-2.61E-09		1.01E-08	-2.68E-09		1.00E-08
Cons	-2.29E-01	***	0.028384	-5.32997	***	0.25738	-5.3328	***	0.255881
Sample	6394			6394			6394		
F(7, 183)	261.73								
Prob>F	0								
R-squared	0.4197								
Root MSE	0.3738								
Wald Chi2(7)				911.15					
Log pseudo				-6357.1			-6174.95		
rho							-0.3864		0.1339
<b>Wald test</b>									
Chi2(1)				6.25			6.7058		
Prob>Chi2				0.0124			0.0096		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.14: Return migrants on detailed labor force participation of nonmigrants at or under 50

Detailed labor force participation at or below 50									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	0.09486	***	0.052964	0.49806	*	0.300838	0.491244	*	0.262292
age	0.032015		0.001927	0.266103	***	0.024442	0.266005	***	0.024039
agesqrd	-0.00015	***	3.73E-05	-0.00288	***	0.000369	-0.00288	***	0.000363
perGDP	7.56E-08		7.11E-08	3.13E-07		4.20E-07	3.15E-07		4.20E-07
fisinc	6.13E-08	**	5.59E-08	1.23E-07		3.12E-07	1.22E-07		3.11E-07
fisspd	-7.61E-08	*	5.69E-08	-2.21E-07		3.08E-07	-2.20E-07		3.07E-07
ttlsale	-6.22E-10		1.53E-09	-5.42E-09		7.93E-09	-5.53E-09		7.90E-09
Cons	-2.27E-01	***	0.023045	-4.76001	***	0.303608	-4.75642	***	0.304662
Sample	6414			6414			6414		
F(7, 184)	1309.88								
Prob>F	0								
R-squared	0.6272								
Root MSE	0.3054								
Wald Chi2(7)				1017.61					
Log pseudo				-5723.85			-5540.49		
rho							-0.25		0.1612
<b>Wald test</b>									
Chi2(1)				1.8			2.207		
Prob>Chi2				0.1802			0.1374		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.15: Return migrants on general employment of nonmigrants at or under 50

Employment under 50									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	0.176257	***	0.067345	0.744658	***	0.25814	0.727561	***	0.226422
age	0.021729	***	0.002021	0.278712	***	0.022035	0.279014	***	0.021433
agesqrd	-4E-05		3.74E-05	-0.00328	***	0.000312	-0.00329	***	0.000304
perGDP	1.50E-07		9.55E-08	5.42E-07		3.73E-07	5.45E-07		3.76E-07
fisinc	9.13E-08		7.70E-08	2.49E-07		3.07E-07	2.51E-07		3.04E-07
fisspd	-6.31E-08		6.64E-08	-1.31E-07		2.61E-07	-1.32E-07		2.59E-07
ttlsale	2.22E-10		3.55E-09	9.31E-10		1.30E-08	8.49E-10		1.29E-08
Cons	-2.16E-01	***	0.029289	-5.44573	***	0.27885	-5.44389	***	0.277504
Sample	6394			6394			6394		
F(7, 183)	223.6								
Prob>F	0								
R-squared	0.4164								
Root MSE	0.3728								
Wald Chi2(7)				860.71					
Log pseudo				-6340.28			-6157.63		
rho							-0.4003		0.137
<b>Wald test</b>									
Chi2(1)				6.02			6.7583		
Prob>Chi2				0.0142			0.0093		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.16: Return migrants on detailed employment of nonmigrants at or under 50

Detailed Employment at or under 50									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	0.100614	*	0.054546	0.525321	*	0.303068	0.527364	**	0.263303
age	0.029478	***	0.001912	0.256593	***	0.024811	0.256162	***	0.024406
agesqrd	-0.0001	***	3.70E-05	-0.00272	***	0.000372	-0.00271	***	0.000367
perGDP	1.03E-07		7.40E-08	4.62E-07		4.36E-07	4.65E-07		4.37E-07
fisinc	4.37E-08		6.00E-08	3.47E-08		3.27E-07	3.18E-08		3.25E-07
fisspd	-6.98E-08		5.72E-08	-1.87E-07		3.04E-07	-1.84E-07		3.03E-07
ttlsale	3.58E-10		1.63E-09	3.41E-10		8.55E-09	1.68E-10		8.53E-09
Cons	-2.18E-01	***	0.023675	-4.72794	***	0.307732	-4.72308	***	0.309303
Sample	6414			6414			6414		
F(7, 184)	1243.93								
Prob>F	0								
R-squared	0.6247								
Root MSE	0.3065								
Wald Chi2(7)				1000.88					
Log pseudo				-5741.27			-5557.66		
rho							-0.2662		0.1619
<b>Wald test</b>									
Chi2(1)				1.91			2.4524		
Prob>Chi2				0.1674			0.1173		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.17: Return migrants on non-agricultural self-employment of nonmigrants at or under 50

Non-agricultural self-employment at or under 50									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.07578	***	0.030589	-0.63268	***	0.237211	-0.58748	***	0.220334
age	0.002926	***	0.000876	0.138885	***	0.022356	0.138634	***	0.022306
agesqrd	3.96E-06		1.71E-05	-0.00166	***	0.00033	-0.00165	***	0.00033
perGDP	2.01E-08		4.78E-08	4.72E-08		2.59E-07	3.68E-08		2.70E-07
fisinc	6.73E-08	*	3.95E-08	4.56E-07	*	2.56E-07	4.56E-07	*	2.58E-07
fisspd	-4.13E-08		3.27E-08	-2.72E-07		2.22E-07	-2.72E-07		2.24E-07
ttlsale	-1.98E-09	**	1.01E-09	-1.63E-08	*	8.54E-09	-1.61E-08	*	8.51E-09
Cons	0.009609		0.012987	-3.7415	***	0.390178	-3.75395	***	0.385212
Size	6414			6414			6414		
F(7, 184)	31.08								
Prob>F	0								
R-squared	0.0263								
Root MSE	0.24125								
Wald Chi2(7)				179.49					
Log pseudo				-5233.89			-5051.22		
rho							0.4532		0.1357
<b>Wald test</b>									
Chi2(1)				9.23			8.1947		
Prob>Chi2				0.0024			0.0042		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.18: Return migrants on being homemakers of nonmigrants at or under 50

Homemaker at or under 50									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.09943	*	0.053894	-0.63949	**	0.304137	-0.62219	**	0.295805
age	0.0108	***	0.001529	0.227623	***	0.02331	0.227373	***	0.023346
agesqrd	-9.99E-05	***	2.73E-05	-0.00301	***	0.00033	-0.003	***	0.00033
perGDP	-1.47E-07	**	6.99E-08	-1.10E-06	*	6.12E-07	-1.11E-06	*	6.12E-07
fisinc	-1.10E-08		5.66E-08	-7.12E-08		3.30E-07	-7.52E-08		3.31E-07
fisspd	-1.45E-08		5.11E-08	-5.38E-08		2.92E-07	-5.02E-08		2.93E-07
ttlsale	4.62E-10		2.35E-09	2.48E-09		1.34E-08	2.51E-09		1.33E-08
Cons	-0.02664		0.022249	-4.60981	***	0.436994	-4.61345	***	0.439354
Sample	6414			6414			6414		
F(7,184)	24.08								
Prob>F	0								
R-squared	0.0779								
Root MSE	0.31177								
Wald Chi2(7)				168.18					
Log pseudo				-5860.32			-5677.51		
rho							0.2833		0.1814
<b>Wald test</b>									
Chi2(1)				2.4			2.1802		
Prob>Chi2				0.1217			0.1398		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.19: Return migrants on full-time study of nonmigrants at or under 30

Full-time Study at 30 or younger									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.15522	***	0.058184	-0.78313	***	0.23197	-0.66037	***	0.197603
age	0.153844	***	0.002859	0.681414	***	0.04133	0.697783	***	0.034849
agesqrd	-0.00559	***	9.66E-05	-0.02565	***	0.001606	-0.02628	***	0.00136
perGDP	-1.09E-07		7.91E-08	-2.87E-07		2.85E-07	-2.75E-07		3.06E-07
fisinc	-9.81E-08	*	5.93E-08	-4.09E-07		2.74E-07	-4.66E-07	*	2.80E-07
fisspd	1.02E-07	**	5.14E-08	4.83E-07	*	2.63E-07	5.38E-07	**	2.67E-07
ttlsale	1.27E-10		1.43E-09	1.12E-09		6.77E-09	8.04E-10		6.72E-09
Cons	-6.33E-02	**	0.028967	-2.57056	***	0.258488	-2.68985	***	0.215697
Sample	3551			3551			3551		
F(7,174)	517.14								
Prob>F	0								
R-squared	0.5098								
Root MSE	0.34356								
Wald Chi2(7)				1022.53					
Log pseudo				-3375.26			-3269.58		
rho							0.4298		0.1112
<b>Wald test</b>									
Chi2(1)				11.1			11.1983		
Prob>Chi2				0.0009			0.0008		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.20: Return migrants on labor force participation of male nonmigrants

Male labor force participation									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
rmhh	0.134535	*	0.072752	0.659305	***	0.264805	0.648958	**	0.211854
age	0.049579	***	0.001226	0.220961	***	0.011924	0.221141	***	0.011336
agesqrd	-0.00053	***	1.59E-05	-0.00234	***	0.000136	-0.00234	***	0.000129
perGDP	1.03E-07		1.02E-07	3.20E-07		3.80E-07	3.21E-07		3.76E-07
fisinc	5.42E-08		7.29E-08	1.14E-07		2.95E-07	1.13E-07		2.93E-07
fisspd	-4.61E-08		6.84E-08	-9.06E-08		2.80E-07	-8.92E-08		2.80E-07
ttlsale	-3.89E-11		2.71E-09	-2.82E-09		9.99E-09	-3.13E-09		9.89E-09
Cons	-3.97E-01	***	0.037115	-4.36503	***	0.176791	-4.35922	***	0.178041
Sample	4230			4230			4230		
F(7, 184)	353.51								
Prob>F	0								
R-squared	0.4267								
Root MSE	0.37798								
Wald Chi2(7)				813.3					
Log pseudo				-4055.47			-3953.23		
rho							-0.4642		0.1257
<b>Wald test</b>									
Chi2(1)				7.59			9.8399		
Prob>Chi2				0.0059			0.0071		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.21: Return migrants on general employment of male nonmigrants

Male employment									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
rmhh	0.139621	*	0.073384	0.693484	***	0.261027	0.680954	***	0.198051
age	0.048432	***	0.001296	0.220621	***	0.012182	0.220849	***	0.011285
agesqrd	-0.00051	***	1.67E-05	-0.00231	***	0.000136	-0.00231	***	0.000127
perGDP	1.48E-07		1.16E-07	5.21E-07		4.42E-07	5.19E-07		4.38E-07
fisinc	3.99E-08		7.63E-08	4.89E-08		3.00E-07	4.86E-08		2.96E-07
fisspd	-4.51E-08		6.79E-08	-8.53E-08		2.70E-07	-8.48E-08		2.68E-07
ttlsale	9.09E-10		3.26E-09	1.66E-09		1.22E-08	1.37E-09		1.22E-08
Cons	-4.00E-01	***	0.037457	-4.48231	***	0.185935	-4.47591	***	0.18541
Sample	4230			4230			4230		
F(7, 184)	310.95								
Prob>F	0								
R-squared	0.4169								
Root MSE	0.38059								
Wald Chi2(7)				818.91					
Log pseudo				-4055.15			-3952.44		
rho							-0.4886		0.1162
<b>Wald test</b>									
Chi2(1)				8.59			12.2495		
Prob>Chi2				0.0034			0.0005		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.22: Return migrants on non-agricultural self-employment of male nonmigrants

Non-agricultural self-employment of male							
Variables	2SLS		IV Probit			Recursive	
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.06933 *	0.037852	-0.49151 **	0.253442	-0.48267 **	0.216953	
age	0.010726 ***	0.000894	0.133976 ***	0.014607	0.133416 ***	0.01469	
agesqrd	-0.00013 ***	1.15E-05	-0.00152 ***	0.00018	-0.00151 ***	0.000181	
perGDP	2.94E-08	4.22E-08	7.25E-08	2.44E-07	6.71E-08	2.41E-07	
fisinc	1.18E-07 ***	4.40E-08	7.19E-07 ***	2.63E-07	7.22E-07 ***	2.64E-07	
fisspd	-6.73E-08 *	3.57E-08	-3.74E-07	2.2E-07	-3.78E-07 *	2.20E-07	
ttlsale	-4.95E-09 ***	1.40E-09	-5.39E-08	2.08E-08	-5.34E-08 ***	2.04E-08	
Cons	-0.04308 **	0.017498	-3.6014 ***	0.311996	-3.59003 ***	0.307811	
Size	4250		4250		4250		
F(7, 184)	41.66						
Prob>F	0						
R-squared	0.052						
Root MSE	0.2721						
Wald Chi2(7)			167.94				
Log pseudo rho			-3488.19		-3386.06		
					0.328	0.1323	
<b>Wald test</b>							
Chi2(1)			4.3		5.2777		
Prob>Chi2			0.0381		0.0216		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.23: Return migrants on being homemakers of male nonmigrants

Male homemaker							
Variables	2SLS		IV Probit			Recursive	
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.10639 *	0.058756	-0.71069 *	0.380264	-0.66722 **	0.333775	
age	0.002149 ***	0.000836	0.062627 ***	0.009162	0.062189 ***	0.009366	
agesqrd	7.44E-06	1.04E-05	-0.00043 ***	0.000084	-0.00043 ***	8.54E-05	
perGDP	-7.70E-08	5.90E-08	-6.65E-07	5.77E-07	-6.62E-07	5.64E-07	
fisinc	-6.44E-09	5.91E-08	-2.12E-08	3.63E-07	-2.29E-08	3.63E-07	
fisspd	1.03E-08	5.17E-08	5.86E-08	3.18E-07	6.04E-08	3.18E-07	
ttlsale	3.27E-09 *	1.85E-09	1.48E-08	9.52E-09	1.47E-08	9.61E-09	
Cons	0.01681	0.025432	-2.98003 ***	0.397375	-2.98383 ***	0.388206	
Sample	4250		4250		4250		
F(7,184)	25.49						
Prob>F	0						
R-squared	0.0353						
Root MSE	0.2768						
Wald Chi2(7)			182.42				
Log pseudo rho			-3519.74		-3418.05		
					0.4432	0.2044	
<b>Wald test</b>							
Chi2(1)			3.69		3.5038		
Prob>Chi2			0.0549		0.0612		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.24: Return migrants on full-time study of male nonmigrants

Male full-time study									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.0986	*	0.059168	-1.04955	***	0.277119	-0.81054	***	0.228183
age	-0.01764	***	0.001399	0.627512	***	0.054629	0.65908	***	0.043878
agesqrd	7.06E-05	***	1.60E-05	-0.02354	***	0.002036	-0.02467	***	0.001656
perGDP	-2.27E-07	***	9.18E-08	-9.06E-07	***	3.65E-07	-9.38E-07	**	4.00E-07
fisinc	-1.08E-07	*	5.94E-08	-1.65E-07		3.33E-07	-2.57E-07		3.50E-07
fisspd	9.18E-08	*	5.15E-08	1.80E-07		3.06E-07	2.52E-07		3.23E-07
ttlsale	4.19E-10		3.61E-09	2.87E-08	*	1.63E-08	3.11E-08	*	1.63E-08
Cons	8.23E-01	***	0.042814	-2.1691	***	0.342838	-2.40165	***	0.281398
Sample	4230			4230			4230		
F(7,184)	198.28								
Prob>F	0								
R-squared	0.4219								
Root MSE	0.33637								
Wald Chi2(7)				670.44					
Log pseudo				-2984.8			-2884.02		
rho							0.5105		0.1256
<b>Wald test</b>									
Chi2(1)				12.05			10.9965		
Prob>Chi2				0.0005			0.0009		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.25: Return migrants on non-agricultural self-employment of female nonmigrants

Non-agricultural self-employment of female									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.05669	**	0.023402	-0.60482	***	0.226353	-0.53093	***	0.194851
age	0.004886	***	0.000444	0.127034	***	0.012284	0.127508	***	0.012108
agesqrd	-5.9E-05	***	5.63E-06	-0.00149	***	0.000146	-0.00149	***	0.000144
perGDP	1.60E-08		5.29E-08	7.19E-08		4.8E-07	6.02E-08		4.89E-07
fisinc	6.40E-08	***	2.40E-08	6.00E-07	***	2.15E-07	6.00E-07	***	2.18E-07
fisspd	-4.91E-08	**	2.17E-08	-4.64E-07	**	2.06E-07	-4.65E-07	**	2.08E-07
ttlsale	-6.40E-10		6.78E-10	-7.55E-09		7.58E-09	-7.32E-09		7.60E-09
Cons	-0.00715		0.011132	-3.7281	***	0.330532	-3.77243	***	0.311132
Size	5864			5864			5864		
F(7, 185)	23.52								
Prob>F	0								
R-squared	.								
Root MSE	0.20828								
Wald Chi2(7)				156.52					
Log pseudo				-4584.36			-4420		
rho							0.4262		0.1227
<b>Wald test</b>									
Chi2(1)				9.44			9.2244		
Prob>Chi2				0.0021			0.0024		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.26: Return migrants on being homemakers of female nonmigrants

Female homemaker									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.1374	*	0.078947	-0.49387	*	0.281222	-0.49742	*	0.271562
age	0.013191	***	0.001323	0.083173	***	0.006167	0.082975	***	0.006217
agesqrd	-1.12E-04	***	1.61E-05	-0.00075	***	6.69E-05	-0.00075	***	6.71E-05
perGDP	-2.75E-07	*	1.57E-07	-1.12E-06		7.58E-07	-1.12E-06		7.55E-07
fisinc	1.21E-07		7.58E-08	4.45E-07	*	2.73E-07	4.47E-07	*	2.73E-07
fisspd	-9.97E-08		6.98E-08	-3.71E-07		2.49E-07	-3.73E-07		2.49E-07
ttlsale	3.16E-09		2.22E-09	1.14E-08		7.40E-09	1.14E-08		7.42E-09
Cons	-0.01843		0.037767	-2.5096	***	0.224496	-2.50288	***	0.226463
Sample	5864			5864			5864		
F(7,185)	43.77								
Prob>F	0								
R-squared	0.0581								
Root MSE	0.4025								
Wald Chi2(7)				396.02					
Log pseudo				-6387.08			-6222.41		
rho							0.2864		0.1638
<b>Wald test</b>									
Chi2(1)				2.69			2.7262		
Prob>Chi2				0.1011			0.0987		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 1.27: Return migrants on full-time study of female nonmigrants

Female full-time study									
Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
rmhh	-0.0708	**	0.036876	-0.71961	***	0.276333	-0.60683	***	0.224323
age	-0.02508	***	0.001137	0.696446	***	0.055554	0.708316	***	0.048376
agesqrd	0.000182	***	1.12E-05	-0.02656	***	0.002118	-0.027	***	0.001855
perGDP	7.57E-08		7.30E-08	1.12E-06		9.30E-07	1.11E-06		9.57E-07
fisinc	-9.47E-08	***	3.44E-08	-6.24E-07	*	3.70E-07	-6.68E-07	*	3.71E-07
fisspd	9.33E-08	***	3.04E-08	7.43E-07	**	3.51E-07	7.88E-07	**	3.48E-07
ttlsale	3.30E-10		1.15E-09	-2.29E-08	***	7.84E-09	-2.33E-08	***	7.86E-09
Cons	8.18E-01	***	0.035312	-2.69748	***	0.346451	-2.79658	***	0.290023
Sample	5826			5826			5826		
F(7,185)	212.85								
Prob>F	0								
R-squared	0.4107								
Root MSE	0.28965								
Wald Chi2(7)				578.35					
Log pseudo				-4134.94			-3971.96		
rho							0.3969		0.136
<b>Wald test</b>									
Chi2(1)				6.16			6.77		
Prob>Chi2				0.013			0.0093		

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.



## CHAPTER 2

### Risk Hedging of Trading Companies in the U.S. Markets

To hedge or not to hedge, that is the question.

#### 2.1 Introduction

This paper attempts to answer the question: Is the use of financial derivatives effective in reducing risks from exchange rate fluctuations for trading companies in the U.S. markets? Here, we do not consider the use of financial derivatives. Firms engaging in import and export businesses confront currency fluctuations as a result of their trading structures and management strategies in addition to using financial derivatives.

In this paper, we study empirically hedging decisions for firms that face foreign exchange-rate risk by first regressing each company's stock returns on returns from several market indexes to obtain the exchange rate exposure of each company. We then regress the predicted exchange rate exposure from the first regression on variables such as the amount of foreign sales out of total sales, level of risk aversion, and level of hedging from each company. The main regression result is that the proportion of foreign sales of a company affects its exchange rate exposure while hedging strategies and risk aversions don't. Companies in the service sector and with total assets under a certain amount are found to be more effective while using their hedging strategies.

The rest of the paper consists of the following sections: section 2.2 summarizes the attitudes toward risk hedging and models used in the risk-hedging literature, section 2.3 describes data scraped, section 2.4 depicts hedging at the firm level using theoretical models, section 2.5 introduces empirical models utilized to analyze the

data, section 2.6 reports the results of the empirical analysis with explanations, section 2.7 checks the robustness of the regression results from section 2.6, and section 2.8 concludes.

## 2.2 Literature Review

Hedging strategies of companies vary based on expectations of losses from foreign exchange rate fluctuations as well as other factors. Allayannis, Lel, and Miller (2003), and Lel (2006) found that management teams of companies under strong external monitoring and governance do a better job in hedging currency risks than their counterparts under weak external monitoring and governance. They used sample firms from 35 and 30 countries respectively between 1990 and 1999. Shareholders also played a role in the making of hedging decisions. Aretz and Bartram (2010), and Hutson and Stevenson (2010) concluded that the wealth of shareholders can be prevented from potential losses from exchange rate fluctuations through hedging. According to Aretz and Bartram (2010):

“direct and indirect costs of financial distress, costly external financing, and taxes, capital market imperfections; economic and legal environment of the country a firm is located in may also affect the decision to hedge.”

Nguyen and Faff (2002) pointed out in their paper that Australian firms did a better job in hedging in the short term than in the long term. Ameer (2010) shared the findings from research on Malaysian companies that volumes of foreign sales and growth opportunities are two major factors when firms consider using financial derivatives for hedging. Managers are incentivized to hedge either with no foreign sales and higher managers' shareholdings or with foreign sales and lower managers' shareholdings but with pressures from institutional investors. Allayannis and Ofek

(2001) used data of nonfinancial firms from S&P 500 in 1993 and drew their conclusion that the proportion of foreign sales in the trade is the only variable that determines the level of derivatives used. This paper, however, utilizes data scraping to parse data from 10-k filings of SEC EDGAR which covers the maximum numbers of U.S. companies of the publicly available data. Two continuous variables measuring risk aversion and hedging strategies respectively using textual analysis have been used on the analysis. Telser (1955) derived from his models the hedging strategies to be used based on the relationships between the expected spot price and the basis. Collins (1997) asserted that it is usually arbitragers and merchandisers who hedge fully while processors hedge partially and farmers seldom hedge at all. Both Telser (1955) and Collins (1997) constructed models in which businesses maximize their expected income subject to the constraints that the probability that their income falls below a certain disaster level should not exceed a certain level. Dellas and Zilberfarb (1993) deduced that risk-averse companies will be able to get rid of exchange rate exposures by selling forward all of their receivables from exports at the expected future price, otherwise, only part of the risks can be hedged through partial hedging.

### **2.3 Descriptive Statistics**

Most of the variables in the data set have been scraped using R. Data from a total of 475 companies have been scraped which reported values in foreign sales and total sales in 10-k filings. Foreign sales, total sales, risk aversion, hedging strategies were parsed from 10-k filings of EDGAR.<sup>1</sup> Foreign sales and total sales were parsed

---

<sup>1</sup>EDGAR is the Electronic Data Gathering, Analysis, and Retrieval system and it performs automated collection, validation, indexing, acceptance, and forwarding of submissions by companies and others.

from the geographical segment of 10-k filings. Textual analysis for risk aversion and hedging strategies has been used to collect as many observations as possible from 10-k filings.<sup>2</sup>

Risk aversion is continuous from 0 to 1, with 0 being risk-loving, and 1 being risk-averse. Numbers of occurrences of each of 17 risk-related keywords<sup>3</sup> were collected from 10-k filings. Kravet and Muslu (2013) searched for risk-related keywords from chapters 1 to 14 of 10-k filing while this paper searched the entire 10-k filing to get the level of risk aversion of each company. The tone of risk aversion is consistent throughout the entire filing of 10-k and all the information related was collected for analysis. Huang et al. (2013) used the difference between numbers of positive words and negative words divided by the total word count to get the tone of a company, while Kravet and Muslu (2013) use the difference between risk-related keywords in the current period and previous period to find the new risk disclosure of a company. This paper aims to find out the level of risk aversion of each company by counting the occurrences of keywords and to make the results comparable among companies. The levels of risk aversion derived from keyword counts are used for analyzing companies' responses to foreign exchange rate fluctuations. Instead of comparing the total word count of risk-related keywords from the 10-k filing of each company, the level of risk aversion has been calculated using mean and median of occurrences of risk-related keywords<sup>4</sup> of each company, which is a value between 0

---

<sup>2</sup>Item 7A., which is Quantitative and Qualitative disclosures about market risk does not apply to "small reporting company" as per Rule 12b-2 of the Exchange Act and it will not be until July 1, 2020, when hedging disclosure will be required from "small reporting company".

<sup>3</sup>Kravet and Muslu (2013), Textual risk disclosures and investors' risk perceptions. This paper excludes the keyword "hedge" from the keyword list since another variable of risk hedging was parsed using keyword searching as well.

<sup>4</sup> $1 - \frac{|Mean_{Risk_i} - Median_{Risk_i}|}{Mean_{Risk_i}}$ , where  $Mean_{Risk_i}$  is the mean value of occurrences of risk-

and  $1 - \frac{|Mean_{Risk_i} - Median_{Risk_i}|}{Mean_{Risk_i}}$  will be larger if word counts of outliers are much larger than other keywords. The outliers (keywords with much higher occurrences) are not as representative of the risk aversion as the keywords that are more closely clustered around the mean of the keyword distribution. The level of risk aversion has been obtained by subtracting the above result from 1.<sup>5</sup> The closer the distribution of the keyword count of a company is to the normal distribution, the more likely the company depicts the market risk more systematically, therefore, its value of risk aversion will be closer to 1. Filings of large corporations tend to be longer and are more likely to have more total keyword count than small companies. The method applied in this paper avoids this issue and makes the results of levels of risk aversion more appropriate for comparison.

Based on keywords from Risk Factors and Quantitative and Qualitative Disclosure About Market Risk as well as other chapters that mention exchange rate risk and hedging strategies from 10-k filings, a total of 32 keywords were selected. There are 10, 18, and 4 keywords for non-hedging, partial hedging, and full hedging respectively. A keyword can be parsed from multiple chapters of a 10-k filing, and the highest word count of occurrences of a keyword among chapters will be used for further analysis. Keywords of one, both, or all three hedging strategies occur from the scraping results. Most companies hedge partially, some companies keep some of their businesses unhedged, partially hedged, or fully hedged. By summarizing the highest occurrences of each keyword from each 10-k filing and grouping those keywords by hedging strategies, the number of occurrences out of the maximum possible occurrences of each hedging group has been calculated. Both total occurrences of

---

related keywords from company  $i$ ,  $Median_{Risk_i}$  is the median value of occurrences of risk-related keywords from company  $i$ .

<sup>5</sup>From 0 to 1, 0 is risk-loving, and 1 is risk-averse.

each strategy and the mean deviation from the median are taken into consideration in creating levels of hedging strategies.  $\left[ \frac{Occurrence_{hi}}{N_h} + \left( 1 - \frac{|Mean_{hi} - Median_{hi}|}{Mean_{hi}} \right) \right]$  has been used to calculate value of partial hedging,  $\left[ \frac{Occurrence_{hi}}{N_h} + \left( 1 - \frac{2|Mean_{hi} - Median_{hi}|}{Mean_{hi}} \right) \right] + 1$  has been used to calculate the value of full hedging, where  $\frac{Occurrence_{hi}}{N_h}$  is the occurrences of total numbers of keywords of a hedging strategy over total numbers of keywords defined for the strategy,  $h=1, 2, 3$ , which stands for the three hedging strategies respectively,  $i$  stands for company. Non-hedging equals 0, the domain of partial hedging is between 0 and 1, and the domain of full hedging is between 1 and 2.  $\frac{Occurrence_{hi}}{N_h}$  has been included because there are occasions when the number of occurrences of two or more keywords is the same, which bring the deviation between the mean and median of occurrences of specific strategies to zero. The value of the hedging strategy of a company is the average of all the values of hedging strategies parsed from the 10-k filing. For example, if there are keywords parsed that belong to non-hedging and partial hedging respectively, the value of the hedging strategy of a company is the average of the values of both non-hedging and partial hedging.

Monthly returns from 2017 to 2019 have been scraped from Yahoo Finance. Prices adjusted for splits and dividends at the first and last trading days of each month of each company have been parsed and are used for calculating the monthly returns. Monthly returns of S&P 500 from 2017 to 2019 are used as the market returns, and monthly returns of Trade-Weighted U.S. Dollar Index: Broad, Goods and Services from FRED are used as the returns of Foreign Exchange Index.

SIC code and Ticker Symbol have been scraped from the SEC website.<sup>6</sup> Ticker symbols are used to scrape monthly prices from Yahoo Finance.

---

<sup>6</sup>Standard Industrial Classification (SIC) are four-digit codes that categorize the industries that companies belong to based on their business activities. A ticker symbol or stock symbol is an abbreviation used to uniquely identify publicly traded shares of a particular stock on a particular stock market. Data were parsed from <https://sec.report/CIK/>.

Table 2.1 summarizes variables used in the empirical analysis. Capital return, market return, and return of FXI are calculated monthly. Data have been parsed from 2017 to 2019 which covers a total of 36 months. Monthly returns of stock prices are used as the proxies for capital returns, market returns are the monthly returns of S&P 500, returns of FXI are the monthly returns of Trade-Weighted U.S. Dollar Index: Broad, Goods and Services. Foreign currency exposure is the predicted value from regression of the capital return on market return and return of FXI. It measures for a 1% increase in the return of FXI, what happens to the capital return<sup>7</sup> of each company. As table 2.1 demonstrates, the mean of the expected exchange rate exposure is -0.09 which means that a 1% increase in the returns of FXI<sup>8</sup> causes stock returns to decrease by 0.09%. Monthly data are selected because according to Allayannis and Ofek (2001): “daily and weekly exchange rate indices are noisier and usually suffer from non-synchronicity problem.”

According to SIC codes, out of 475 companies, 50.53% are in manufacturing followed by 27.37% which are in services. Finance,<sup>9</sup> insurance, and real estate ranks third (7.29%), and transportation, communications, electric, gas, and sanitary service ranks fourth (6.88%). Detailed classification of sectors can be found in the appendix in table 2.2 and table 2.3.

Table 2.4 summarizes the distribution of hedging strategies by sector. Table 2.5 depicts the distribution of total assets.

---

<sup>7</sup>Proxied by stock returns.

<sup>8</sup>The US dollar depreciates.

<sup>9</sup>Companies that are either bank holding or investment banks are excluded from regression analysis.

Table 2.1: Descriptive Statistics

	Mean	Std	Q3	Median	Q1
Capital Return	0.009982	0.031694	0.023476	0.011008	-0.01589
Market Return	0.009406	0.035712	0.02775	0.01445	-0.0261
Return of FXI	-0.0013	0.014153	0.011905	-0.00182	-0.01955
Exchange Rate Exposure	-0.09337	1.667989	0.6742	-0.0587	-1.7938
FS/TS	0.3728	0.247727	0.52502	0.34409	0.0713
Risk Aversion	0.4782	0.122484	0.5502	0.4666	0.3243
Hedge Strategy	0.5655	0.314379	0.6674	0.5882	0
Total Asset	1.61E+12	5.12E+12	9.88E+11	1.51E+11	3.08E+08

## 2.4 The Theoretical Model

Collins (1997) created a positive model of hedging which was based on the fundamental idea of Telser's (1955) return maximization model and Turvey and Baker's (1990) interpretation of hedging as a financial decision. The model explains the reasons behind various hedging decisions<sup>10</sup> based on financial structure and risk aversion.

Each company being confronted with foreign exchange-rate fluctuations maximizes expected equity subject to the constraint that the probability of terminal equity drops below certain level is less than a prespecified value. The equity maximization problem can be written as:

$$\max \bar{E}_{j1} = \int_{-\infty}^{\infty} E_{j1} g(E_{j1}) dE_{j1} \quad (j = 1, \dots, n) \quad (2.1)$$

$$s.t. \int_{-\infty}^{jd} g(E_{j1}) dE_{j1} \leq j\alpha \quad (j = 1, \dots, n) \quad (2.2)$$

Where  $g(E_{j1})$  is the probability density function for terminal equity of company  $j$ .

It is a two-period model which starts at period 0 and ends at period 1.

---

<sup>10</sup>Non-hedging, partial hedging, and full hedging.



Mathematically, a company  $j$  maximizes its expected wealth at the end of the period 1 given that the possibility of expected wealth falls below  $j\alpha$ .

The expected wealth of company  $j$  is

$$\bar{E}_{j1} = E_{j0} + [p_{jh}H_j + \bar{p}_{jc}(1 - H_j)]Y_j - (V_{jh}H_jY_j) - k_jY_j - i_jD_j - F_j \quad (2.3)$$

where  $E_{j0}$  is initial equity of  $j$ ,  $E_{j1}$  is end-of-period equity of  $j$ ,  $p_{jh}$  is the net price of hedge contracts for  $j$  (net of commission, margin interest, etc.),  $\bar{p}_{jc}$  is random cash price to be received on unhedged output for  $j$ ,  $H_j$  is hedge ratio of  $j$ ,  $Y_j$  is output of  $j$  in units,  $V_{jh}$  is the cost of hedge contracts,  $k_j$  is variable cost of production per unit of  $Y_j$ ,  $i_j$  is interest rate paid by  $j$  on debt,  $D_j$  is the debt of  $j$ ,  $F_j$  is total fixed costs of production of  $j$ .

To study the impact of the extent of hedging on terminal equity,

$$\frac{\partial \bar{E}_{j1}}{\partial H_j} = (p_{jh} - \bar{p}_{jc} - V_{jh})Y_j \quad (2.4)$$

The auxiliary function can be defined as:

$$L_1 = \bar{E}_{j1} - \lambda_{j1} \left[ \int_{-\infty}^{j\alpha} g(E_{j1}) dE_{j1} - j\alpha \right] - \lambda_{j2} (V_{jh} - p_{jh} - \bar{p}_{jc})Y_j, \quad \text{if } \frac{\partial \bar{E}_{j1}}{\partial H_j} \geq 0 \quad (2.5)$$

To maximize  $\bar{E}_{j1}$  by selecting  $H_j$ , the second constraint will not be binding, because  $(p_{jh} - \bar{p}_{jc} - V_{jh})Y_j \neq 0$ , therefore,  $\lambda_{j2} = 0$ .<sup>11</sup>

If  $p_{jh} > \bar{p}_{jc} + V_{jh}$ , the terminal equity will increase with the increase of the proportion of output hedged according to Equation 2.4.<sup>12</sup> It is assumed that the probability that the terminal equity falls below certain level also decreases with the increase in  $H$ . Since companies hedge for risk reduction, and not for speculation, the incentive to hedge is when there is the probability that the terminal equity will

---

<sup>11</sup>complementary slackness

<sup>12</sup> $Y_j$  is nonnegative.

drop below zero, which is the bankruptcy level.<sup>13</sup> Equation 2.2 can be rewritten as

$$\int_{-\infty}^0 g(E_{j1}) dE_{j1} = 0 \quad (j = 1, \dots, n) \quad (2.6)$$

Equation 2.6 is the new budget constraint, and Equation 2.5 can be rewritten as:

$$L_2 = \bar{E}_{j1} - \gamma_{1j} \left[ \int_{-\infty}^0 g(E_{j1}) dE_{j1} \right] - \gamma_{j2} (V_{jh} - p_{jh} - \bar{p}_{jc}) Y_j \quad (2.7)$$

When  $H=0$ , which means that no proportion of the output will be hedged, and equation 2.3 becomes:

$$\bar{E}_{j1} = E_{j0} + \bar{p}_{jc} Y_j - k_j Y_j - i_j D_j - F_j \quad (2.8)$$

Assuming that equity and cost of a company are exogenous, and it is cash price and debt that determine the hedge decision. Substitute the worst possible terminal equity  $E_{ja}$  and the worst possible cash price  $p_{ja}$  into equation 2.8:

$$\bar{E}_{ja} = E_{j0} + p_{ja} Y_j - k_j Y_j - i_j D_j - F_j \quad (2.9)$$

$E_{ja}$  should be nonnegative for the company to stay away from bankruptcy.

$$\bar{E}_{ja} \geq 0 \quad (2.10)$$

From equation 2.9 and 2.10:

$$p_{ja} \geq \frac{k_j Y_j + i_j D_j + F_j - E_{j0}}{Y_j} \quad (2.11)$$

which means that a company will choose not to hedge if the worst possible cash price is greater than the financial obligations per unit of output.  $Y_j = \gamma_j A_j = \gamma_j (D_j + E_{j0})$ ,

---

<sup>13</sup>This is the most conservative assumption. A company can choose to hedge as much as possible.

where  $\gamma_j$  is the Leontief coefficient of units of output per dollar of assets in company  $j$ , and  $A_j$  is the total asset of company  $j$ . The financial obligation of a company  $j$  is:

$$\theta_j = \frac{k_j Y_j - i_j D_j - F_j - E_{j0}}{\gamma_j (D_j + E_{j0})} \quad (2.12)$$

$$\frac{d\theta_j}{dD_j} = \frac{(k_j \gamma_j + i_j) \gamma_j (D_j + E_{j0}) - \gamma_j (k_j Y_j + i_j D_j + F_j - E_{j0})}{[\gamma_j (D_j + E_{j0})]^2} = \frac{\gamma_j [E_{j0}(1 + i_j) - F_j]}{[\gamma_j (D_j + E_{j0})]^2} \quad (2.13)$$

$Y_j$  is a function of  $D_j$  and  $E_{j0}$ . The result is positive so long as  $E_{j0}(1 + i_j) > F_j$ . Financial obligation increases with increase in its debt. When financial leverage increases or when cash price decreases to such an extent that Equation 2.11 is violated, which results into  $\bar{E}_{ja}=0$ , then the company will start to hedge.  $\bar{E}_{ja}$  now equals to

$$\bar{E}_{ja} = E_{j0} + [p_{jh}H_j + \bar{p}_{ja}(1 - H_j)]Y_j - (V_{jh}H_jY_j) - k_jY_j - i_jD_j - F_j \quad (2.14)$$

By dividing both sides of the Equation 2.14 by  $Y_j$

$$\frac{\bar{E}_{ja}}{Y_j} = (\bar{p}_{ja} - k_j) + (p_{jh} - \bar{p}_{ja} - V_{jh})H_j + \frac{E_{j0} - i_jD_j - F_j}{Y_j} = 0 \quad (2.15)$$

By substituting  $Y_j$  with  $\gamma_j (D_j + E_{j0})$ , the optimal proportion to hedge is

$$H_j^* = \frac{(k_j - \bar{p}_{ja}) + \frac{i_j D_j + F_j - E_{j0}}{\gamma_j (D_j + E_{j0})}}{p_{jh} - \bar{p}_{ja} - V_{jh}} = \frac{D_j (\gamma_j k_j - \gamma_j \bar{p}_{ja} + i_j) + E_{j0} (\gamma_j k_j - \gamma_j \bar{p}_{ja} - 1) + F_j}{\gamma_j (D_j + E_{j0}) (p_{jh} - \bar{p}_{ja} - V_{jh})} \quad (2.16)$$

When there is a change in financial leverage:

$$\frac{\partial H_j^*}{\partial D_j} = \frac{i_j \gamma_j (D_j + E_{j0}) - \gamma_j (F_j + i_j D_j - E_{j0})}{(p_{jh} - \bar{p}_{ja} - V_{jh}) [\gamma_j (D_j + E_{j0})]^2} = \frac{\gamma_j [(1 + i_j) E_{j0} - F_j]}{(p_{jh} - \bar{p}_{ja} - V_{jh}) [\gamma_j (D_j + E_{j0})]^2} \quad (2.17)$$

The result is positive if  $(1 + i_j) E_{j0} > F_j$ , and  $p_{jh} > \bar{p}_{ja} + V_{jh}$ .<sup>14</sup>

---

<sup>14</sup>Random cash price decreases, and net price of hedge contracts surpasses the cash price.

When there is a change in cash price:

$$\frac{\partial H_j^*}{\partial \bar{p}_{ja}} = \frac{(-1)(p_{jh} - p_{ja} - V_{jh}) - (-1) \left[ (k_j - \bar{p}_{ja}) + \frac{i_j D_j + F_j - E_{j0}}{\gamma_j (D_j + E_{j0})} \right]}{(p_{jh} - p_{ja} - V_{jh})^2} \quad (2.18)$$

By dividing both the numerator and denominator by  $p_{jh} - p_{ja} - V_{jh}$  and rearranging the terms, there goes

$$\frac{\partial H_j^*}{\partial \bar{p}_{ja}} = \frac{H_j^* - 1}{p_{jh} - \bar{p}_{ja} - V_{jh}} \quad (2.19)$$

which is nonpositive because  $0 \leq H^* \leq 1$  and  $p_{jh} - p_{ja} - V_{jh} > 0$ .

To sum up, when  $p_{jh} > \bar{p}_{ja} + V_{jh}$ , since terminal equity is an increasing function of the proportion to hedge, a company will choose to hedge as much proportion of its output as possible as long as the cost of the hedging contracts remains the same.<sup>15</sup> Ceteris paribus, a company will choose to hedge 100% of its equity. An increase in financial obligation and cash price will cause the optimal proportion to hedge to increase and to decrease respectively.<sup>16</sup>

When  $p_{jh} < \bar{p}_{ja} + V_{jh}$ , according to Equation 2.4, an increase in the proportion to hedge will decrease the terminal equity. To follow Equation 2.6, if the terminal equity is nonnegative,  $H_j = 0$ ; when the terminal equity drops below zero, signing of hedge contracts will decrease the terminal equity instead of increasing it. From Equations 2.16 and 2.17, if  $(1 + i_j) E_{j0} > F_j$ , and  $p_{jh} < \bar{p}_{ja} + V_{jh}$ , An increase in financial obligation and cash price will cause the optimal proportion to hedge to decrease and to increase respectively. Under this scenario, the company will not hedge. Depending on financial structure and level of risk aversion, the upper bound of the budget constraint and the threshold varies.

---

<sup>15</sup>Marginal calls from futures contract and a possible increase in cash price will cause the favorable condition to be reversed.

<sup>16</sup>Based on the conditions from Equations 2.16 and 2.17.

## 2.5 The Empirical Model

This paper analyzes how U.S. companies engaging in foreign trade respond to foreign exchange-rate fluctuations. There are two preliminary questions to answer for the analysis to be carried out: how to measure foreign exchange rate fluctuations confronted by each company and how to measure hedging strategies taken by each company to mitigate foreign exchange rate risk. Following Jorion (1990), Amihud (1993), and Allayannis (1996), the following empirical model can be applied to analyze the exchange rate exposure of a company:

$$R_{it} = \beta_{0i} + \beta_{1i}R_{mt} + \beta_{2i}FXI_t + \epsilon_{it}, t = 1, \dots, T \quad (2.20)$$

where

$R_{it}$  is the rate of return on the  $i^{th}$  firm's common stock in period  $t$ ,

$R_{mt}$  is the rate of return of the market portfolio in period  $t$ ,

$FXI_t$  is the rate of return of the trade-weighted exchange rate index, measured in US dollar per unit of foreign currencies in period  $t$ .

This regression is designed to obtain the expected value of  $\beta_{2i}$ , which is interpreted as: for a 1% change in the rate of return of the foreign exchange-rate index, what will be the impact on the rate of return of a company's common stock. This expected value is used as the exchange rate exposure of a company.  $\hat{\beta}_{2i}$  can be interpreted the same way as the market beta in the Capital Asset Pricing Model.<sup>17</sup> Market beta measures the riskiness of an investment as compared to the market portfolio, while  $\hat{\beta}_{2i}$  measures the responsiveness of capital returns of a company against the change in the trade-weighted exchange rate index, controlling for mar-

---

<sup>17</sup> $ER_i = R_f + \beta_i(ER_m - R_f)$ , where  $ER_i$  is the expected return of investment,  $R_f$  is the risk free rate,  $(ER_m - R_f)$  is the market risk premium, and  $\beta_i$  is the market beta.

ket returns. Allayannis and Ofek (2001) assume that the difference between nominal and real exchange rates is small and that therefore there is little difference in using risk premiums in this context. The impact of the exchange rate on the rate of return of a company outweighs the impact of inflation. Variations in the risk-free rate are rather negligible compared to variations in exchange rates.

Three years of monthly data have been collected and one exchange rate exposure has been obtained from Equation 2.20 for each company. Monthly returns of S&P 500 are used as proxies of monthly market returns and monthly returns of Trade weighted U.S. Dollar Index from FRED are used as monthly returns of Foreign Exchange Index. Wider time frames from 2015 to 2019 are taken into consideration. 2020 hasn't been included because of the pandemic which had a great impact on international trade.

Equation 2.21 measures given risk aversion, what are the impacts of hedging strategies on exchange rate exposure? Comparing between partial hedging and non-hedging, and full hedging and non-hedging respectively, any difference will be made on hedging against exchange rate exposure?

$$\hat{\beta}_{2i} = \alpha_{0i} + \alpha_{1i}FS/TS_i + \alpha_{2i}Risk_i + \alpha_{3i}Partial\_Hedge_i + \alpha_{4i}Fully\_Hedge_i + \eta_i \quad (2.21)$$

Where

$\hat{\beta}_{2i}$  is the estimated exchange rate exposure from Equation 20,

$FS/TS_i$  is the ratio of foreign sales out of total sales of company i,

$Risk_i$  is a continuous value representing risk level of company i, with 0 being risk loving, and 1 being risk averse,

$Partial\_Hedge_i$  and  $Fully\_Hedge_i$  are a continuous value representing hedging strategies of company i, the domains are from 0 to 1 and 1 to 2 respectively.

## 2.6 Regression Results

FXI from Equation 2.20 is the rate of return of trade-weighted exchange rate index denominated in US dollar per unit of foreign currency. If the rate of return of FXI increases, i.e., the trade-weighted exchange rate index increases, which means that one unit of foreign currency can exchange for more US dollars, therefore, the US dollar depreciates against foreign currencies. An increase in FXI indicates depreciation of the US dollar. For exporters, if there are revenues to be collected in foreign currencies, exporters benefit from depreciation when exchanging foreign currencies for the US dollar. Exchange rate exposure<sup>18</sup> decreases following an increase of the FXI. Importers suffer from the depreciation of the US dollar because they need to spend more US dollars for the exchange of foreign currencies to purchase goods and services overseas. The impact of exchange-rate fluctuations on a company depends on the overall exposures from exporting and importing businesses if a company is both an exporter and an importer. A company's exchange rate exposure is influenced by its exporting business only if the company is only an exporter. A company's exchange rate exposure is influenced by its importing business only if the company is only an importer. When the US dollar depreciates, which means FXI increases, and if the overall performance of a company from foreign trade suffers from a loss, stock returns drop as well, and vice versa.<sup>19</sup> Under this condition, FXI and stock returns are negatively correlated. When the US dollar appreciates, importers can take advantage of the appreciation, and FXI and stock returns are positively correlated. The impact of currency fluctuations on a foreign

---

<sup>18</sup>In particular, translation exposure: the risk that a company's equities, assets, liabilities, or income will change in value as a result of exchange rate changes.

<sup>19</sup>When the US dollar depreciates, possible cases that cause losses are: trade deficit for exporting and importing companies, companies doing importing businesses.

trade company can be ambiguous. For example, dollar appreciation will cause the export to decrease because of more expensive dollars for foreign importers, while it is beneficial for importers. If a company imports and exports, the effect of dollar appreciation can be uncertain. From table 2.1, the mean of exchange rate exposure is -0.09337. Generally speaking, the rate of return in FXI is negatively correlated with stock returns.

Tables 2.6 and 2.7 use exchange rate exposure as the dependent variable which has been derived from 3 years of data. Regression results from regressions 1 to 6 show a negative correlation between FS/TS and exchange rate exposure and are significant at over 96% confidence intervals. The coefficient of the variable of interest is positive and insignificant from regression 7 and 8. Risk aversions, hedging strategies, and sectors are statistically insignificant. To avoid collinearity, variables of non-hedging strategy and the mining sector are not included in the regression. Regression 1 and 2 do not control for sectors. Regression 1 tells the difference between partial hedging and full hedging, while regression 2 combines partial hedging and full hedging as one hedging strategy. From regression 1 to 8, odd-numbered regressions tell the difference between partial hedging and full hedging, while even-numbered regressions don't. Regression 1 indicates a 0.75% decrease in exchange rate exposure following a 1% increase in FS/TS, and it is statistically significant at 97% confidence interval. Regression results from regression 2 don't vary much. When controlling for sectors, the results drop slightly and are significant at about 96% confidence interval. When restricting samples to manufacturing, according to regression 5, a 1% increase in FS/TS results in a 0.98% decrease in exchange rate exposure, and it is significant at 97% confidence interval. The manufacturing sector shows significant result because it involves more frequent exchange of goods and is subject to changes in costs of capital internationally.



Results from regressions 7 and 8 are positive between exchange rate exposures and hedging strategies. For a company from the service sector, on average, a 1 unit increase in the level of partial hedging increases exchange rate exposure by 1.61, and a 1 unit increase in the level of hedging strategies in general increases exchange rate exposure by 1.15. Both coefficients are significant at about 94% confidence intervals. Hedging intensifies the negative correlation between stock returns and the return of FXI in the service sector. When FXI increases, USD depreciates, stock return decreases, and hedging accelerates the decrease in stock returns. When FXI decreases, USD appreciates, stock return increases, and hedging accelerates the increase in stock returns. Hedging acts as an indicator of a company's performance as a result of exchange-rate fluctuations. When the US dollar depreciates, and a company hedges against losses, investors' prediction of the company's performance becomes less optimistic which is to the disadvantage of stock returns. When the US dollar appreciates, the company hedges to prevent potential losses in the future, and investors are more optimistic about the company's performance which is to the advantage of stock returns.

Ceteris paribus, an increase in FS/TS results in the decrease of exchange rate exposure which means an increase in FS/TS weakens the negative relationship between the rate of return of FXI and stock returns.<sup>20</sup> When the return of FXI increases, the US dollar depreciates which renders an increase in exports, thus, FS/TS increases. Although stock return decreases following the depreciation of the US dollar, the increase in exports counteracts some loss from translation exposure. When the return of FXI decreases, the US dollar appreciates which causes a decrease in exports, therefore, FS/TS decreases. Although stock return increases because of

---

<sup>20</sup>Exchange rate exposure has been derived from the coefficient of the relationship between returns FXI and returns of common stocks.

the appreciation, the decrease in exports counteracts some of the gains from negative translation exposure. Both risk aversion and hedge strategies are statistically insignificant. Neither risk aversion nor hedging strategies show any significant correlations with exchange rate exposure. For a company that is risk-averse and when the company predicts increasing exchange rate exposure, it will reduce its transactions in international trade, or to remain its current transactions and hedge against the increased risks, therefore, hedging strategies play a supplementary role in risk-averse companies. For a company that is risk-loving, its transactions in international trade will not be influenced greatly by exchange-rate fluctuations, and its use of hedging strategies doesn't fully reflect the exchange rate exposures that the company is being confronted with.

In sum, FS/TS has a direct impact on a company's exchange rate exposure, while risk aversion and hedging strategies are uncorrelated with exchange rate exposure in general. The level below which a company starts to hedge varies, therefore, regression of sample data doesn't show any significant correlations between exchange rate exposure and hedging strategies. Debt and cash flows are not included since those two variables are correlated with the extent of hedging instead of the hedging decisions to be made.

## **2.7 Robustness Checks**

Regression results are based on 475 non-financial companies. The predicted exchange rate exposure for each company is obtained from analyzing three years of data on stock returns, market returns, and FXI from 2017 to 2019. To check the

regression results, two more years of data from 2015 to 2016<sup>21</sup> have been added to come up with new predicted exchange rate exposures. Empirical equations used on regressions 9 to 16 (Tables 2.8 and 2.9) are the same as used on regressions 1 to 8, and the only difference is that the dependent variable is the predicted exchange rate exposure derived from 5 years of market data instead of 3. Regression results still show a negative correlation between exchange rate exposure and FS/TS. The coefficient of the main variable of interest from regression 9 is -0.41, which is the result without controlling for sectors or differentiating between partial hedging and full hedging. It is significant at 91% confidence interval with a decrease in magnitude compared to the coefficient of -0.75 with a significance level of 97% from regression 2.1. Such data of market returns in 2015 and 2016 have a weaker correlation with corporate data of 2019 than data of market returns from 2017 to 2019. Regression results from 10 to 16 show negative correlations between exchange rate exposures and FS/TS with decreased coefficients in magnitudes and significance levels. From regression results 11 and 12, compared to the mining sector, construction, manufacturing, transportation and communication, etc., wholesale, retail, and services are more susceptible to exchange rate exposure.

Regression results from Table 2.10 to 2.13 control for total assets. Coefficients of FS/TS on exchange rate exposure remain at the same level.

Among the 475 companies, 237 report total assets that are above 151 billion US dollars, which is about 50% of the data set. Companies with total assets above 151 billion US dollars don't show significant correlations between exchange rate exposure and FS/TS, nor are there any other significant correlations. 238 companies reported total assets that are at or below 151 billion. Tables 2.14 and 2.15 show

---

<sup>21</sup>International trade in 2020 has undergone stagnation because of the pandemic and was therefore excluded.

regression results of companies with total assets under 151 billion and exchange rate exposures are derived from regression results from 3 years and 5 years of data respectively. The result from regression 33 in Table 14 shows that a 1% increase in FS/TS decreases exchange rate exposure by about 1.09% and it is significant at 98% confidence intervals. Companies with relatively lower total assets experience a more efficient reduction in exchange rate exposure comparing to companies with higher total assets.

## 2.8 Conclusion

Based on 475 corporate data which are scraped from SEC EDGAR, and exchange rate exposure derived from 3 years and 5 years of market data respectively,<sup>22</sup> regression results show negative correlations between exchange rate exposure and FS/TS. Risk aversion and hedging strategies are uncorrelated with exchange rate exposures in general. Companies in the service sector show positive correlations between hedging strategies and exchange rate exposure.

Hedging doesn't play a prominent role in dealing with exchange-rate fluctuations among U.S. companies in general in international trade. The proportion of foreign sales out of total sales has a direct impact on exchange rate exposure, and those companies with higher export would benefit from hedging.

Furthermore, hedging strategies can be used prudently for companies in the service sector since the effect on capital returns can be quite the opposite under dollar appreciation and depreciation. Companies with total assets under a certain level are more efficient in reducing exchange rate exposure.

---

<sup>22</sup>Regressions were firstly run using exchange rate exposure derived from 3 years of market data, then using 5 years of market data.

## 2.9 Appendix

Table 2.2: Sectors

Industry	Count	Percentage
<b>Mining</b>	11	2.32%
Metal Mining	1	0.21%
Coal Mining	0	0.00%
Oil and gas extraction	9	1.88%
Mining and quarrying of nonmetallic minerals, except fuels	1	0.21%
<b>Construction</b>	7	1.47%
Building construction-general contractors and operative builders	2	0.42%
Heavy construction other than building construction-contractors	3	0.63%
Construction-special trade contractors	2	0.42%
<b>Manufacturing</b>	240	50.53%
Food and kindred products	13	2.71%
Tobacco products	0	0.00%
Textile mill products	3	0.63%
Apparel and other finished products made from fabrics and similar material	3	0.63%
Lumber and wood products, except furniture	2	0.42%
Furniture and Fixtures	1	0.21%
Paper and allied products	3	0.63%
Printing, publishing, and allied industries	3	0.63%
Chemicals and allied products	52	10.83%
Petroleum refining and related industries	0	0.00%
Rubber and miscellaneous plastics products	3	0.63%
Leather and leather products	1	0.21%
Stone, clay, and concrete products	2	0.42%
Primary metal industries	8	1.67%
Fabricated metal products, except machinery and transportation equipment	9	1.88%
Industrial and commercial machinery and computer equipment	35	7.29%
Electronic and other electrical equipments, except computer	37	7.71%
Transportation equipment	59	12.29%
Miscellaneous manufacturing industries	6	1.25%
<b>Transportation, communications, electric, gas, and sanitary service</b>	33	6.95%
Railroad transportation	1	0.21%
Motor freight transportation and warehousing	4	0.83%
United States Postal Service	0	0.00%
Water Transportation	2	0.42%
Transportation by air	3	0.63%
Pipelines, except natural gas	1	0.21%
Transportation services	3	0.63%
Communications	11	2.29%
Electric, gas, and sanitary services	8	1.67%
<b>Wholesale trade</b>	13	2.74%
Wholesale trade: durable goods	9	1.88%
Wholesale trade: nondurable goods	4	0.83%

Table 2.3: Sector continues

<b>Retail trade</b>	11	2.32%
Building materials, hardware, garden supply, and mobile home dealers	2	0.42%
Food stores	0	0.00%
Automotive dealers and gasoline service stations	0	0.00%
Apparel and accessory stores	1	0.21%
Home furniture, furnishing, and equipment stores	1	0.21%
Eating and drinking places	3	0.63%
Miscellaneous retail	4	0.83%
<b>Finance, insurance, and real estate</b>	30	6.32%
Depository institutions	4	0.83%
Nondepository credit institutions	2	0.42%
Security and commodity brokers, dealers, exchange, and services	11	2.29%
Insurance carriers	6	1.25%
Real estate	2	0.42%
Holding and other investment offices	10	2.08%
Industry	Count	Percentage
Services	129	27.16%
Hotels, rooming houses, camps, and other lodging places	0	0.00%
Personal services	3	0.63%
Business services	100	20.83%
Automotive repair, services, and parking	0	0.00%
Motion pictures	4	0.83%
Amusement and recreation services	4	0.83%
Health services	8	1.67%
Legal services	3	0.63%
Social services	0	0.00%
Museums, art galleries, botanical, and zoological gardens	0	0.00%
Membership organizations	0	0.00%
Engineering, accounting, research, management, and related services	7	1.46%
Private household	0	0.00%
<b>Public administration</b>	1	0.21%
Services, not elsewhere classified	1	0.21%
Executive, legislative, and general government, except finance	0	0.00%
Justice, public order, and safety	0	0.00%
Public finance, taxation, and monetary policy	0	0.00%
Administration of human resource programs	0	0.00%
Administration of environmental quality programs	0	0.00%
Administration of economic programs	0	0.00%
National Security and international affairs	0	0.00%
Nonclassifiable establishments	0	0.00%

Table 2.4: Distribution of hedging strategies by sector

Industry	NonHdg	ParHdg	FulHdg	NA
Mining	0	11	0	0
Construction	0	6	1	0
Manufacturing	16	210	14	0
Transportation	2	31	0	0
Wholesale	1	12	0	0
Retail	0	11	0	0
Fin Service	0	29	1	0
Service	12	113	5	0

Table 2.5: Distribution of total asset

Total Asset above	Count	Percentage
1.51E+11	238	50.11%
1.00E+11	250	52.63%
5.00E+10	267	56.21%
1.00E+10	322	67.79%
5.00E+09	353	74.32%
3.00E+09	368	77.47%
2.00E+09	390	82.11%
1.00E+09	406	85.47%
5.00E+08	422	88.84%
1.00E+08	442	93.05%
659219	475	100%

Table 2.6: Regression results: all sectors, 3 years

Variables	Regression 1		Regression 2		Regression 3		Regression 4	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.7486	0.0256**	-0.7695	0.0213**	-0.678	0.0474**	-0.6966	0.0408**
Risk	-0.4132	0.5291	-0.4234	0.5187	-0.3478	0.6014	-0.3558	0.5926
ParHdg	0.2033	0.5851			0.3187	0.4007		
FulHdg	0.5268	0.1404			0.5925	0.1020		
ParNFul			0.3763	0.2156			0.4663	0.1323
Cons					0.0046	0.9956	0.0167	0.9839
Manu					-0.013	0.9808	0.0024	0.9964
Trans					-0.0907	0.8840	-0.0895	0.8856
WhSal					0.1753	0.8062	0.1814	0.7995
Retal					0.3455	0.6447	0.3594	0.6311
FinSvs					0.303	0.6221	0.3157	0.6071
Service					0.3602	0.5142	0.3814	0.4887
Intercept	0.2729	0.4449	0.2216	0.5282	0.0422	0.9484	-0.0169	0.9791
No. of Obs.	419		419		419		419	
$R^2$	0.0075		0.0083		0.0016		0.0030	

FS/TS is the ratio between foreign sales and total sales, Risk is risk aversion with domain from 0 to 1; 0 means risk loving, and 1 means risk averse, ParHdg is partial hedging with domain from 1 to 2, FulHdg means full hedging with domain from 1 to 2, non-hedging has been dropped to avoid collinearity, ParNFul combines both partial and full hedging strategies, Cons means construction, Manu means manufacturing, Trans means transportation, WhSal means wholesale, Retal means Retail, FinSvs means financial services.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.7: Regression results: manufacturing and service, 3 years

Variables	Regression 5		Regression 6		Regression 7		Regression 8	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.9753	0.0319**	-0.9894	0.0300**	-0.0355	0.9670	0.0966	0.9095
Risk	-0.2658	0.7881	-0.4345	0.6600	-0.2984	0.8290	-0.2451	0.8592
ParHdg	-0.2597	0.6208			1.606	0.0600*		
FulHdg	0.5524	0.2393			0.8326	0.2700		
ParNFul			0.2298	0.5860			1.1544	0.0742*
Intercept	0.3700	0.4782	0.2677	0.6060	-0.3167	0.6820	-0.2365	0.7576
No. of Obs.	210		210		115		115	
$R^2$	0.0173		0.0104		0.0007		0.0036	

Regression 5 and 6 are from manufacturing , and regression 7 and 8 are from service.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.



Table 2.8: Regression results: all sectors, 5 years

Variables	Regression 9		Regression 10		Regression 11		Regression 12	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.4126	0.0923*	-0.4281	0.0796*	-0.3989	0.1060*	-0.4110	0.0944*
Risk	-0.1235	0.7970	-0.1303	0.7860	0.0070	0.9884	0.0018	0.9970
ParHdg	0.0835	0.7579			0.1446	0.5953		
FulHdg	0.3154	0.2209			0.3154	0.2210		
ParNFul			0.2092	0.3423			0.2379	0.2828
Cons					1.5374	0.0090***	1.5446	0.0086***
Manu					1.2487	0.0010***	1.2584	0.0011***
Trans					1.3936	0.0020***	1.3944	0.0017***
WhSal					1.0831	0.0331**	1.0869	0.0324**
Retal					1.5539	0.0040***	1.5626	0.0034***
FinSvs					1.2279	0.0050***	1.2350	0.0052***
Service					1.4459	0.0003***	1.4595	0.0002***
Intercept	0.0621	0.8142	0.0244	0.9252	-1.3178	0.0050***	-1.3551	0.0035***
No. of Obs.	403		403		403		403	
$R^2$	0.0006		0.0015		0.0210		0.0226	

FS/TS is the ratio between foreign sales and total sales, Risk is risk aversion with domain from 0 to 1; 0 means risk loving, and 1 means risk averse, ParHdg is partial hedging with domain from 1 to 2, FulHdg means full hedging with domain from 1 to 2, non-hedging has been dropped to avoid collinearity, ParNFul combines both partial and full hedging strategies, Cons means construction, Manu means manufacturing, Trans means transportation, WhSal means wholesale, Retal means Retail, FinSvs means financial services.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.9: Regression results: manufacturing and service, 5 years

Variables	Regression 13		Regression 14		Regression 15		Regression 16	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.1648	0.5870	-0.1699	0.5760	-0.6874	0.3420	-0.5428	0.4420
Risk	-0.1386	0.8360	-0.2159	0.7450	0.1723	0.8770	0.2272	0.8380
ParHdg	-0.0998	0.7760			1.1823	0.0850*		
FulHdg	0.2604	0.4040			0.4693	0.4280		
ParNFul			0.1191	0.6720			0.7530	0.1430
Intercept	0.0150	0.9660	-0.0303	0.9310	-0.2408	0.7040	-0.1731	0.7840
No. of Obs.	204		204		107		107	
$R^2$	-0.0118		-0.0122		-0.0054		-0.0046	

Regression 13 and 14 are from manufacturing , and regression 15 and 16 are from service.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.10: Regression results: all industry, 3 years controlling TA

Variables	Regression 17		Regression 18		Regression 19		Regression 20	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.7562	0.0247**	-0.7772	0.0205**	-0.6831	0.0464**	-0.7017	0.04**
Risk	-0.4232	0.5201	-0.4336	0.5096	-0.3547	0.5949	-0.3627	0.5860
ParHdg	0.2026	0.5867			0.3177	0.4028		
FulHdg	0.5253	0.1404			0.5916	0.1030		
ParNFul			0.3751	0.2175			0.4653	0.1340
Total Assets	-4.9E-15	0.7447	-5E-15	0.7407	-3.7E-15	0.8103	-3.6E-15	0.8110
Cons					0.0046	0.9956	0.0167	0.9840
Manu					-0.0122	0.9821	0.0033	0.9950
Trans					-0.0814	0.8962	-0.0802	0.8980
WhSal					0.1795	0.8019	0.1856	0.7950
Retal					0.3443	0.6463	0.3582	0.6330
FinSvs					0.3055	0.6196	0.3182	0.6050
Service					0.3603	0.5145	0.3816	0.4890
Intercept	-0.2894	0.423	-0.2385	0.00621	0.0530	0.9355	-0.0063	0.9920
No. of Obs.	419		419		419		419	
$R^2$	0.0053		0.0062		-0.0007		0.0007	

FS/TS is the ratio between foreign sales and total sales, Risk is risk aversion with domain from 0 to 1; 0 means risk loving, and 1 means risk averse, ParHdg is partial hedging with domain from 1 to 2, FulHdg means full hedging with domain from 1 to 2, non-hedging has been dropped to avoid collinearity, ParNFul combines both partial and full hedging strategies, Cons means construction, Manu means manufacturing, Trans means transportation, WhSal means wholesale, Retal means Retail, FinSvs means financial services.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.11: Regression results: manufacturing and service, 3 years controlling TA

Variables	Regression 21		Regression 22		Regression 23		Regression 24	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.9871	0.0306**	-1.002	0.0287**	-0.1609	0.8545	-0.0089	0.9917
Risk	-0.2830	0.7754	-0.4521	0.6478	-0.3515	0.8	-0.2901	0.8340
ParHdg	-0.2566	0.6258			1.669	0.0515**		
FulHdg	0.5531	0.2379			0.8267	0.2741		
ParNFul			-0.2315	0.5841			1.1760	0.0696*
Total Assets	-7.27E-15	0.6986	-7.63E-15	0.6858	5.05E-14	0.3404	4.69E-14	0.3737
Intercept	0.3929	0.4553	0.2920	0.5771	-0.3465	0.6548	-0.2576	0.7372
No. of Obs.	210		210		115		115	
$R^2$	0.0132		0.0064		-0.0001		0.0018	

Regression 21 and 22 are from manufacturing , and regression 23 and 24 are from service.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.12: Regression results: all industries, 3 years controlling TA

Variables	Regression 25		Regression 26		Regression 27		Regression 28	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.4246	0.0844**	-0.4401	0.0726**	-0.4102	0.0976*	-0.4224	0.0867*
Risk	-0.1395	0.7719	-0.1464	0.7608	-0.0080	0.9868	-0.0131	0.9782
ParHdg	0.0822	0.7617			0.1421	0.6019		
FulHdg	0.3129	0.2250			0.3133	0.2244		
ParNFul			0.2072	0.3473			0.2357	0.2877
Total Asset	-6.84E-15	0.5303	-6.89E-15	0.5268	-7.00E-15	0.5187	-6.99E-15	0.5192
Cons					1.537	0.0090***	1.5450	0.0086***
Manu					1.251	0.0013***	1.2600	0.0011***
Trans					1.411	0.0015***	1.4120	0.0015***
WhSal					1.091	0.0321**	1.0950	0.0313**
Retal					1.551	0.0037***	1.5600	0.0035***
FinSvs					1.234	0.0053***	1.2410	0.0050***
Service					1.446	0.0003***	1.4600	0.0002***
Intercept	0.0872	0.7447	0.0499	0.8498	-1.296	0.0058***	-1.3330	0.0042***
No. of Obs.	403		403		403		403	
R <sup>2</sup>	-0.0009		1.56E-05		0.0195		0.0212	

FS/TS is the ratio between foreign sales and total sales, Risk is risk aversion with domain from 0 to 1; 0 means risk loving, and 1 means risk averse, ParHdg is partial hedging with domain from 1 to 2, FulHdg means full hedging with domain from 1 to 2, non-hedging has been dropped to avoid collinearity, ParNFul combines both partial and full hedging strategies, Cons means construction, Manu means manufacturing, Trans means transportation, WhSal means wholesale, Retal means Retail, FinSvs means financial services.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.13: Regression results: manufacturing and service, 5 years controlling TA

Variables	Regression 29		Regression 30		Regression 31		Regression 32	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.1852	0.5430	-0.1905	0.5320	-0.7124	0.3297	-0.5592	0.4330
Risk	-0.1704	0.7990	-0.2473	0.7100	0.1577	0.8881	0.2171	0.8460
ParHdg	-0.0951	0.7870			1.199	0.0831*		
FulHdg	0.2609	0.4040			0.4683	0.4314		
ParNFul			0.1213	0.6660			0.7575	0.1420
Total Asset	-1.21E-14	0.3310	-1.23E-14	0.3250	1.30E-14	0.7567	9.94E-15	0.8120
Intercept	0.0558	0.8760	0.0117	0.9740	-0.2506	0.6945	-0.1794	0.7770
No. of Obs.	204		204		107		107	
R <sup>2</sup>	-0.0121		-0.0124		-0.0144		-0.0139	

Regression 29 and 30 are from manufacturing , and regression 31 and 32 are from service.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.14: Regression results: total assets under 151 billion, 3 years

Variables	Regression 33		Regression 34		Regression 35		Regression 36	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-1.0930	0.0247**	-1.1150	0.0215**	-1.0750	0.0319*	-1.0980	0.0278**
Risk	-0.0249	0.9799	-0.0181	0.9854	0.1834	0.8553	0.1873	0.8519
ParHdg	0.5718	0.2644			0.6059	0.2454		
FulHdg	0.9098	0.0748*			0.9245	0.0776*		
ParNFul			0.7424	0.0832*			0.7646	0.0809*
Total Asset	8.93E-12	0.0202**	8.83E-12	0.0213**	9.33E-12	0.0171**	9.25E-12	0.0178**
Cons					0.6080	0.6287	0.6887	0.5809
Manu					0.9535	0.2251	0.9754	0.2133
Trans					1.1370	0.2014	1.1420	0.1989
WhSal					0.9099	0.3766	0.9061	0.3788
Retal					0.9847	0.3622	0.9917	0.358
FinSvs					1.577	0.0775*	1.6020	0.0720*
Service					1.446	0.0003* * *	1.0590	0.1857
Intercept	-0.2970	0.6019	-0.3555	0.5259	1.0360	0.1968	1.5020	0.1219
No. of Obs.	213		213		213		213	
$R^2$	0.0374		0.0403		0.0218		0.0251	

FS/TS is the ratio between foreign sales and total sales, Risk is risk aversion with domain from 0 to 1; 0 means risk loving, and 1 means risk averse, ParHdg is partial hedging with domain from 1 to 2, FulHdg means full hedging with domain from 1 to 2, non-hedging has been dropped to avoid collinearity, ParNFul combines both partial and full hedging strategies, Cons means construction, Manu means manufacturing, Trans means transportation, WhSal means wholesale, Retal means Retail, FinSvs means financial services.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 2.15: Regression results: total assets under 151 billion, 5 years

Variables	Regression 37		Regression 38		Regression 39		Regression 40	
	Estimate	P value	Estimate	P value	Estimate	P value	Estimate	P value
FS/TS	-0.6050	0.1050*	-0.6187	0.0957*	-0.6617	0.0772*	-0.6753	0.0699*
Risk	0.1259	0.8680	0.1327	0.8602	0.3125	0.6747	0.3163	0.6703
ParHdg	0.3419	0.3800			0.4521	0.2416		
FulHdg	0.5560	0.1500			0.6377	0.0984*		
ParNFul			0.4510	0.1656			0.5453	0.0925*
Total Asset	2.29E-12	0.4320	2.23E-12	0.4422	3.40E-12	0.2387	3.36E-12	0.2434
Cons					1.2760	0.1675	1.3230	0.149
Manu					1.9080	0.0011***	1.9210	0.0010***
Trans					2.1930	0.0009***	2.1960	0.0009***
WhSal					1.5480	0.0417**	1.5460	0.0415**
Retal					2.1790	0.0064***	2.1840	0.0062***
FinSvs					1.9820	0.0032***	1.9960	0.0029***
Service					2.1090	0.0004***	2.1230	0.0004***
Intercept	0.2463	0.5800	0.2852	0.5148	-2.3100	0.0017***	-2.3540	0.0012***
No. of Obs.	206		206		206		206	
$R^2$	0.0013		0.0050		0.0404		0.0444	

FS/TS is the ratio between foreign sales and total sales, Risk is risk aversion with domain from 0 to 1; 0 means risk loving, and 1 means risk averse, ParHdg is partial hedging with domain from 1 to 2, FulHdg means full hedging with domain from 1 to 2, non-hedging has been dropped to avoid collinearity, ParNFul combines both partial and full hedging strategies, Cons means construction, Manu means manufacturing, Trans means transportation, WhSal means wholesale, Retal means Retail, FinSvs means financial services.

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

## CHAPTER 3

### Migration in India under Structural Transformation

#### 3.1 Introduction

As per the 2001 census of India, there are 314 million migrants, out of which 268 million (85%) were intra-state migrants (those who migrate from one part of the state to another) and 41 million (13%) were inter-state migrants (those who migrate from one state to another). Migration requires a certain amount of savings from the migrants as well as the ability to deal with the uncertainties of the job market at the destination. What makes migration in India unique is that there are 122 major languages and 1,599 other languages as reported in the 2001 census data of India. Apart from language barriers, there still remains the caste system to some extent, adding another hurdle for the migrants to face. Yet, Mahapatro (2012) points out that the numbers of people who migrated from underdeveloped districts to better-off districts have increased over time as the Indian government improved transportation and other infrastructure making travel between states easier, improving access to information, thereby rendering job search easier. As a result, currently, inter-state migrant job seekers outnumber their intra-state migrant counterparts.

Kone, Liu, Mattoo, Ozden, and Sharma (2017) emphasize the difficulties that Indian inter-state migrants are confronted with in another state: having access to social benefits and tertiary education, as well as receiving job offers from the public sector. In India, only residents holding “ration cards” can get subsidized food and get admission to local public hospitals. The ration cards are only issued and accepted by the home state government. State residents are given priority in admissions to universities and technical institutes. For employment in government entities, state domicile is a prerequisite.

Thus, on the one hand, some improvements facilitate inter-state migration, but on the other, there still remain the force of social norms and policies that restrict inter-state migration. By comparing inter-state migrants with intra-state migrants in terms of labor force participation, this paper attempts to answer the question: do inter-state migrants take a more active role in the job market than their intra-state counterparts? Are there any differences among male and female migrants in terms of employment?

The rest of the paper consists of the following sections: section 3.2 summarizes the topics studied in the literature of internal migration in India, section 3.3 describes the sample and subsamples, section 3.4 introduces the empirical model utilized for analysis as well as the identification strategy, section 3.5 discusses the regression results, section 3.6 tests the robustness of the regression results and the first stage correlation of the instrumental variable, section 3.7 provides the conclusion.

## **3.2 Literature Review**

The literature studying the internal migration in India focuses on regional internal migration, gender differences in internal migration, women in internal migration, social class in internal migration, the role of infrastructure in internal migration, and the role of remittances in internal migration.

Using the 1999 to 2000 National Sample Survey of India, Singh (2009) finds increases in female migrants exceeded increases in male migrants in Mumbai. Regarding the motivation for migration, male migrants were found to be employment-oriented while female migrants were marriage and family-oriented. From 1980 to 2000, male migrants were engaged in jobs in production, while female migrants were commonly seen working in community and social service sectors. As the population

continues to expand, Singh expressed his concern about the lack of policies that benefit both migrants and locals. The issue of internal migrants exists not only in Mumbai but also in other parts of India as well. Agnihorti, Mazumdar, and Neetha (2013) focus on differences between male and female migrants in India. Social hierarchies such as scheduled tribes and scheduled castes restrict female migrants to short-term and circular migration, with some exceptions as female migrants working in textile industries who suffer from no caste-specific bias. Female migrants are more concentrated in the agricultural sector while male migrants are more engaged in the industry and the service sectors. Despite the recent fall in the proportion of female migrants in the agricultural sector, gender bias in the employment of migrants still remains. Kaur (2004) addresses the issue of female migration from the perspective of marriage. With the decreasing sex ratio<sup>1</sup> among teenagers and the asymmetric distribution of marriageable girls across India, it is not uncommon to find a bride and a groom who were from different regions with different cultural backgrounds, with women from poor families marrying men from rich families for upward social mobility. While the number of single men exceeds the number of single women in some districts in India, there exist social norms against inter-caste marriage. Though the Indian government outlawed the caste system in 1948, Bhagat (2009) studies the mobility of Indians from various social classes and finds that scheduled castes and scheduled tribes who are not from any of the four major varnas don't show any higher mobility than Indians from other social classes. Upper-class Indians who are better off were found to be more likely to migrate than the underclass. Apart from social norms, the level of development of both the source and target district has been found to play an important role in the decision to migrate. Asturias, Ramos, and Santana (2019) base their study on the Golden Quadrilateral (GQ) of India.

---

<sup>1</sup>Number of females per 1,000 males.



GQ is the construction of a modern 5,800 km highway that was designed to connect India's four major metropolitan areas (Delhi, Mumbai, Chennai, and Calcutta). Inhabitants living closer to the GQ are found to be more likely to migrate than inhabitants living further away from the GQ. Castaldo, Deshingkar, and McKay (2012) focus on the role of remittances in internal migration and poverty among Indian and Ghanaian migrants respectively. Deshingkar (2006) found that internal remittances are much smaller than international remittances, but households receive internal remittances much more frequently than international remittances. Internal remittances contribute to consumption smoothing, alleviate credit constraints and reduce vulnerability. In what follows, we contribute to the literature of internal migration in India by analyzing inter-state and intra-state migrants and their labor force participation. As it is more costly to migrate out of state, this paper attempts to find out if inter-state migrants secure more job opportunities than their intra-state counterparts.

### 3.3 Descriptive Statistics

The data source of this paper is from the survey data of India of IPUMS<sup>2</sup>International.<sup>3</sup> The survey data of India are available at the household level in 1983, 1987, 1993, 2004, and 2009. There are 100,957 households in the survey of 2009. The following is a summary of natives, inter-state migrants, intra-state migrants in terms of their gender, age, education, labor force participation, and sector of occupation. Table

---

<sup>2</sup>IPUMS stands for Integrated Public Use Microdata Series. University of Minnesota, National Statistical Offices, international data archives, and other international organizations have contributed to the collection, maintenance, and dissemination of the database.

<sup>3</sup>IPUMS-International is a database of census microdata from around the world. There are 473 census and surveys from 102 countries and statistics of over 1 billion people have been recorded so far.

3.1 to 3.4 are summaries of descriptive statistics. The city and state where the surveyee inhabited in 2004 and 2009 respectively were used to define the migration status of the surveyees. If a surveyee stayed in the same state and the same city between 2004 and 2009, then he/she is defined as a native, if a surveyee stayed in the same state but stayed in different cities between 2004 and 2009, then he/she is an intra-state migrant, if a surveyee stayed in different states between 2004 and 2009, then he/she is an inter-state migrant.

Natives make up less than 1% in all the classifications of the data. The largest age group among inter-state migrants is those aged between 35 and 44 which makes up 27.28% of the subsample of inter-state migrants. The age group that ranks the top among intra-state migrants is that aged between 45 and 54 which makes up 27.22% of the subsample of intra-state migrants. There are 28.66% of male migrants aged between 35 and 44 which is the largest group among all the age groups of male migrants followed by 25.41% of male migrants aged between 45 and 54. There are 24.35% of female migrants aged between 45 and 54 which ranks the highest followed by 21.36% of female migrants aged between 55 and 64. Male inter-state migrants who were aged 34 or younger make up 9.88% of the male inter-state subsample. Male intra-state migrants who were 34 or younger make up 18.18% of the male intra-state subsample. Female inter-state migrants who were 34 or younger make up 14.39% of the female inter-state subsample. Female intra-state migrants who were 34 or younger make up 13.81% of the female intra-state subsample. Surveyees between 25 and 54 years of age constitute the largest part of the sample and subsamples by gender and migration types.

Inter-state migrants who were illiterates rank the top in the subsample of inter-state migrants which took 24.95% of the subsample of inter-state migrants followed by 23.37% of inter-state migrants whose highest level of education were primary

school. With a proportion of 21.71%, illiterates of intra-state migrants comprise the largest group among the subsample of the education level of intra-state migrants followed by 20.13% of intra-state migrants whose highest level of education was primary school. The percentage of male surveyees whose highest level of education was primary school is 23.54% which ranks the highest among all the male subsamples of education groups. Male illiterates comprise 21.76% of the subsample which ranks second in the subsample. Female illiterates make up 48.14% of the female subsample of education groups which is the largest group. The second largest group is females whose highest level of education was primary school and they take 21.19% of the subsample. Illiterates and people with only education in primary school make up 45.32% of the male subsample. Illiterates and people with only education in primary school make up 70.44% of the female subsample. Male inter-state migrants whose highest level of education was primary school or lower make up 45.41% of the subsamples of male inter-state migrants. Female inter-state migrants whose highest level of education was primary school or lower make up 70.60% of the subsample of female inter-state migrants. Male intra-state migrants whose highest level of education was primary school or lower make up 38.03% of the subsample of male intra-state migrants. Female intra-state migrants whose highest level of education was primary school or lower make up 66.67% of the subsample of female intra-state migrants. Education levels were low be it by gender or by migration types. As are indicated by the statistics, the percentages of female migrants with a low level of education are much higher than male migrants.

There were 43.84% of inter-state migrants who were self-employed, which is the highest proportion among the subsample of inter-state migrants. There were 41.84% of inter-state migrants who were wage-employed, which is the highest proportion among the subsample of intra-state migrants. 41,252 male inter-state migrants

were self-employed which accounts for 47.05% of the subsample of male inter-state migrants, which was the highest proportion in the subsample. The most prominent group among female inter-state migrants was wage-employment, there were 2,939 of them which makes 25.65% of the subsample of female inter-state migrants. With a total number of 605 which makes 44.16% of the subsample of male intra-state migrants, male intra-state migrants who worked as wage workers ranks at the top on their list of employment types. The highest number of female intra-state migrants who worked as wage-workers is 56 which comprises 26.67% of the subsample of female intra-state migrants by employment types. Migrants play a more active role in self-employment be it among inter-state migrants or male inter-state migrants, while intra-state migrants engaged more in wage-employment and the same thing goes with female subsamples.

The top three sectors where there is the most number of male inter-state migrants are skilled agricultural and fishery workers, elementary occupations, and crafts and related trade workers, which constitutes 20.77%, 18.37%, and 12.09% of the subsample of male inter-state migrants respectively. The top three sectors where there is the most number of male intra-state migrants are elementary occupations, skilled agricultural and fishery worker, and legislators, senior officials, and manager, which constitutes 21.24%, 13.94%, and 10.15% of the subsample of male intra-state migrants respectively. The sectors where the total number of female inter-state migrants ranks the top three were elementary occupations, skilled agricultural and fishery worker, and service workers, and shop and market sales, which accounts for 15.30%, 10.38%, and 5.46% of the subsample of female inter-state migrants respectively. The sectors where the total number of female intra-state migrants ranks the top three were elementary occupations, skilled agricultural and fishery worker, and service workers, and shop and market sales, which takes up 15.71%, 4.76%, and

4.76% of the subsample of female intra-state migrants respectively. The sectors of occupations that are among the top of the list of both male and female inter-state migrants are skilled agricultural and fishery workers, and elementary occupations. The sectors of occupations that are among the top of the list of both male and female intra-state migrants are skilled agricultural and fishery workers, and elementary occupations, which are the same as their inter-state counterparts.

### 3.4 Empirical Models and Identification Strategy

To answer the question of whether inter-state migrants are taking a more active role in labor force participation than their intra-state counterparts, the data set was analyzed using three empirical models with the endogeneity of the migration decision being tackled with an instrumental variable. Empirical models applied in this paper include a two-stage least squares (2SLS), a probit with instrumental variable (IV probit), and a recursive bivariate probit Model (Recursive). The instrumental variable used is the total number of households of inter-state migrants in 2009 over the total number of households of inter-state migrants<sup>4</sup> in history<sup>5</sup> at the district level.<sup>6</sup> The reasoning behind the use of this IV is as follows. Inter-state migrants are confronted with obstacles in language, culture, and policies in the destination districts. With no inter-state migrants in a district, it is difficult for the first few of them to find jobs and to at least temporarily settle down there. If there is already a certain number of inter-state migrants working in the destination district,

---

<sup>4</sup>Hereinafter referred to as “the ratio”.

<sup>5</sup>Data are available in 1987, 1999, 2004, and 2009.

<sup>6</sup>The household itself is excluded from both the numerator and denominator when applying the IV to each household if the household is included as a household with inter-state migrants while creating the variable.

the presence of a network of migrants lowers the barriers to a successful job search. The higher the ratio in a district, the more likely the surveyee will migrate across state borders to that district. The increase in the number of inter-migrants is an indicator that the district is more receptive and there is an increase in the demand for labor from other states.

For 2SLS, we have

$$Y_i^* = a_0 + a_1 M_i^* + a_2 X_i + a_3 D_i + a_4 DT_i + e_i \quad (3.1)$$

$$\hat{M}_i = \alpha_0 + \alpha_1 MRto_i + \alpha_2 X_i + \alpha_3 D_i + \alpha_4 DT_i \mu_i \quad (3.2)$$

$$Y_i^* = \beta_0 + \beta_1 \hat{M}_i + \beta_2 X_i + \beta_3 D_i + \beta_4 DT_i + \epsilon_i \quad (3.3)$$

where

$MRto_i$  denotes the ratio.

$X_i$  denotes personal characteristics.

$D_i$  denotes the district level GDP.

$DT_i$  denotes the driving distance between two districts.

$Y_i^*$  and  $M_i^*$  are unobserved latent variables with

$$Y_i = \begin{cases} 1 & \text{if } Y_i^* > 0 \\ 0 & \text{if } Y_i^* \leq 0 \end{cases} \quad (3.4)$$

and

$$M_i = \begin{cases} 1 & \text{if } M_i^* > 0 \\ 0 & \text{if } M_i^* \leq 0 \end{cases} \quad (3.5)$$

$Y_i^*$  and  $M_i^*$  are expected values from labor force participation and the migration decisions which are obtained from the answers of the questionnaire: if a surveyee is an inter-state migrant, we decide it to be the case that by revealed preference the expected return from being an inter-state migrant outweighs the expected return of

being an intra-state migrant or nonmigrant. A dummy variable of return migrant has been created from the questionnaire, with inter-state migrant equaling 1, and 0 otherwise. A similar logic applies to the expected return to join the workforce and other specific employment types. For example, if a surveyee chose to become a wage worker, this implies that for this surveyee the expected return from being wage-employment exceeds the expected return from other choices of employment types. Thus, in this case, the dependent variable equals 1, otherwise, it equals 0.  $M R t o_i$  is the instrumental variable used to deal with the endogeneity of the migration decision.  $X_i$  denotes personal characteristics, and  $D_i$  denotes the district level GDP of the surveyee measured in crore rupees.<sup>7</sup>  $D T_i$  denotes the driving distance between two districts if the migrant is an inter-state migrant, as scraped from the Google map.  $e_i$ ,  $\mu_i$ , and  $\epsilon_i$  denote the error terms.

The probit model that takes into account the endogeneity issue analyzes the same variables as in the 2SLS model.

The recursive bivariate probit model is as follows:

$$P r o b [ Y = 1, M = 1 | \mathbf{X}_1, \mathbf{X}_2 ] = \phi ( \theta_1 \mathbf{X}'_1 + \lambda M, \theta_2 \mathbf{X}'_2, \delta ) \quad (3.6)$$

where  $\mathbf{X}_1$  and  $\mathbf{X}_2$  are the individual characteristics that influence the migration and labor force participation decisions of migrant workers. According to Greens (1983), “the endogenous nature of one of the variables on the right-hand side of the first equation can be ignored in formulating the log-likelihood.” As can be seen from Equations 3.7 and 3.8, the second regressand,  $M_i$ , appears on the right-hand side of Equation 3.7. This model is recursive and simultaneous. Equation 3.8 analyzes the migration decisions of the migrants and Equation 3.7 analyzes the labor force participation decisions given their migration decisions. The econometric

---

<sup>7</sup>One crore equals to ten million rupee.

specification is as follows:

$$Y_i^* = \gamma_0 + \gamma_1 M_i + \gamma_2 X_i + \gamma_3 D_i + \gamma_4 DT_i + \delta_i \quad (3.7)$$

$$M_i^* = \rho_0 + \rho_1 MRto_i + \rho_2 X_i + \rho_3 D_i + \rho_4 DT_i + \tau_i \quad (3.8)$$

One of the assumptions of the recursive bivariate probit model is that the decisions on labor force participation and migration are interdependent, with  $cov(\delta_i, \tau_i) = \sigma$ . To obtain a more robust result to distributional misspecification, according to Monfardini and Radice (2008), an instrument  $MRto_i$  is used, which is the same IV used in the previous models.

### 3.5 Regression Results

The following three empirical models were utilized in the regression analysis: a two-stage least square model (2SLS), a probit model with instrumental variable (IV probit), and a recursive bivariate probit model (Recursive). Comparisons between the labor force participation of the inter-state migrants and intra-state migrants regarding the following aspects have been made: being in the labor force, being self-employed, being wage-employed, being an employer, being a day laborer, and being an unpaid family worker.

Inter-state migrants show a stronger tendency in labor force participation in general, Furthermore, inter-state migrants are more likely to work as wage-employees or as day laborers. Intra-state migrants, however, show a higher probability of becoming unpaid family workers. The regression results are statistically significant at 90% confidence intervals or higher. Inter-state migrants are found to be more active in being self-employed or taking the position as employers, nevertheless, the results are statistically insignificant.



Table 3.5 shows the regression results with labor force participation as the dependent variable and migration type as the independent variable. Both variables are dichotomous. If a surveyee is in the labor force, the value of the dependent variable equals 1, otherwise, it equals 0. If the surveyee is an inter-state migrant, the value of the independent variable equals 1, otherwise, it equals 0.<sup>8</sup> The coefficient from the regression results of the recursive model indicates that inter-state migrants are more likely to join the labor force than their intra-state counterparts with the magnitude of 0.48, and it is statistically significant at 90% confidence intervals. Regression results from the 2SLS and the IV probit model are consistent with the results from the Recursive model but are statistically insignificant. Inter-state migrants travel longer distances than intra-state migrants to their destinations and are more likely to be confronted with barriers in local cultures and policies than their intra-state counterparts. Although inter-state migrants have to deal with all or some of those disadvantages, they are willing to take the risks and challenges and seek opportunities from the job market in the target state. The regression result is an indicator that barriers for migrants to work in other states have been alleviated and some states provide more job opportunities.

Table 3.6 shows the results from regressing the wage-employment of the surveyees on their migration status. The IV probit model yields a positive correlation between being wage-employed and being inter-state migrants (the magnitude is 2.34). The result is statistically significant at 90% confidence intervals. Regression results from the 2SLS and the Recursive model are statistically insignificant. The regression results in general are consistent with the results as shown in Table 3.5. Inter-state migrants are more likely to participate in the labor force than intra-state migrants

---

<sup>8</sup>Dependent variables and variables of interest in the following regression results are all dichotomous and the values are assigned the same way as described above.

and they are more likely to become wage workers.

Table 3.7 is the regression results on the probability of inter-state migrants to be day laborers.<sup>9</sup> Both the IV probit and Recursive model show positive correlations between being day laborers and being inter-state migrants, and the magnitudes are 2.6 and 1.2 respectively. Regression results from the IV probit and recursive model are statistically significant at 90% and 99% confidence intervals respectively. Regression results from the 2SLS model are positive but insignificant. The regression results between Table 3.7 and Table 3.5 are more consistent than the regression results between Table 3.6 and Table 3.5, especially in the Recursive model. Regression result from Table 3.5 shows positive correlations between being in the labor force and being inter-state migrants, while regression result from Table 3.7 shows positive correlations between being a day laborer and being an inter-state migrant. To work as a day laborer is more in line with the type of work that inter-migrants choose. As shown by the statistics, in 2009, the percentages of male and female inter-state migrants with high school education or lower were 76.98% and 88.43% respectively, while the percentages of male and female intra-state migrants with high school education or lower were 74.31% and 85.71% respectively. In general, migrants don't have a strong educational background, and there are still cultural obstacles as well as local employment policies that are in favor of the locals to be dealt with.

Regression results from Table 3.8 indicate that inter-state migrants are less likely to become unpaid family workers. The magnitude of coefficients from IV probit and recursive models are 4.2 and 2.17 respectively. The coefficients from IV probit and recursive models are statistically significant at 95% and 99% confidence intervals respectively.

---

<sup>9</sup>Day laborers are those who work and get paid daily.

### 3.6 Exogeneity Checks and the Relevance of Instrumental Variable

Regression results from the first stage of the 2SLS and IV probit model and the second stage of the Recursive model are shown in Table 3.9. Coefficients from all of the three regressions indicate that the choice of being an inter-state migrant is positively correlated with the total number of households of inter-state migrants in 2009 over the total number of households of inter-state migrants in history (the ratio). The magnitudes of the coefficients from the regression results of the 2SLS, IV probit, and Recursive model are 0.14, 0.14, and 6.58 respectively. The regression results are statistically significant at 99% confidence intervals from all three models. The higher the ratio of a district, the more likely a potential migrant will travel across state borders to become an inter-state migrant in that state.

Further analysis has been done by checking gender, age, and the combination of both gender and age on the decision making of inter-state migrants of being in the labor force, being a wage worker, being an employer, being a day laborer, and being an unpaid family worker.<sup>10</sup> Summary of the significance and sign of the regressions from 2SLS, IV probit, and Recursive model can be found in Table 3.10.

Description of the regression results from the male subsample is as follows. Male inter-state migrants are more likely to become employers than their intra-state counterparts, however, male inter-state migrants are found to be less likely than male intra-state migrants to be employers in the Recursive model. However, male inter-state migrants who are under 30 are less likely than their intra-state counterparts to be in the labor force as indicated by the regression results from the IV probit and

---

<sup>10</sup>The unpaid family worker is regarded as the group who are temporarily unemployed and looking for jobs.

Recursive model, and both are statistically significant at 99% confidence intervals. However, the same subsample is found to be more likely to become unpaid family workers as indicated by the regression result from the recursive model. In general, male inter-state migrants play a less active role in the job market compared to their intra-state counterparts. Factors in the target city such as dialects, religious beliefs as well as local employment policies are barriers to job search, however, female inter-state migrants who face similar obstacles appear to do better in terms of employment as regressions from the female samples are statistically significant in being in the labor force, being a wage worker, and being a day laborer. Regression results from the female subsample are shown in Tables 3.11 to 3.13. Female inter-state migrants have a higher probability than female intra-state migrants to be in the labor force, and they are more likely to work as wage workers and as day laborers than their intra-state counterparts. Regression results for female inter-state migrants under 30 are statistically significant at 95% confidence intervals from the IV probit model in labor force participation. Regression results are statistically significant at 95% confidence intervals from the IV probit model in wage employment. Regression results are statistically significant at 99% confidence intervals from the Recursive model in being a day laborer. In short, female inter-state migrants appear to have done a better job in the job market than their intra-state as well as inter-state male counterparts.

### **3.7 Conclusion**

Given the time and effort they invest in migration and job search, the incentives of inter-state migrants to find jobs are stronger. As a result, the probability for inter-state migrants to engage in wage employment and to work as day laborers is

higher. Regression results show that female inter-state migrants are more successful than their intra-state counterparts in job market participation in general, and in wage employment and being day laborers specifically. These results stand in sharp contrast to those obtained for the younger than 30 male inter-state migrants, who are less likely to be employed than their intra-state counterparts. An analysis of the reasons for this finding will need to be carried out.

### 3.8 Appendix

Table 3.1: Migration status by age

Migration type	Male				Female			
	Native	Intra-state migrants	Inter-state migrants	Total	Native	Intra-state migrants	Inter-state migrants	Total
<b>Age</b>								
0 to 14	0%	3.23%	96.78%	155	1.08%	3.23%	95.70%	93
15 to 19	0.21%	5.60%	94.19%	482	0.66%	5.30%	94.04%	151
20 to 24	0.05%	1.38%	98.57%	1882	0%	1.17%	98.83%	256
25 to 34	0.15%	1.26%	98.59%	15180	0.67%	1.26%	98.06%	1188
35 to 44	0.18%	1.35%	98.47%	25569	0.37%	1.94%	97.69%	2424
45 to 54	0.25%	1.70%	98.05%	22684	0.28%	1.58%	98.14%	2849
55 to 64	0.30%	1.67%	98.03%	14238	0.32%	1.92%	97.76%	2500
65 +	0.36%	1.69%	97.95%	9054	0.53%	1.82%	97.65%	2252

Table 3.2: Migration status by education

Migration type	Male				Female			
	Native	Intra-state migrants	Inter-state migrants	Total	Native	Intra-state migrants	Inter-state migrants	Total
<b>Education</b>								
Illiterate	0.24%	1.32%	98.44%	19422	0.28%	1.52%	98.20%	5719
Primary Sch	0.31%	1.26%	98.43%	21027	0.28%	2.09%	97.63%	2532
Mdl Sch	0.27%	1.80%	97.93%	15149	0.79%	2.28%	96.93%	1272
Hgh Sch	0.15%	1.71%	98.14%	13086	0.85%	1.33%	97.82%	827
College	0.22%	1.39%	98.39%	7626	0.69%	2.94%	96.37%	579
Grad	0.10%	1.90%	98.00%	12919	0.38%	1.66%	97.95%	782
Unknown	0%	0%	100%	15	0%	0%	100%	2

Table 3.3: Migration status by employment

Migration type	Male			Total	Female			Total
	Native	Intra-state migrants	Inter-state migrants		Native	Intra-state migrants	Inter-state migrants	
<b>Employment</b>								
Self wage	0.19%	1.35%	98.46%	41899	0.22%	0.90%	98.88%	2226
Unpd	0%	0.57%	99.43%	176	0%	1.75%	98.25%	57
Niu	0.45%	2.21%	97.34%	8960	0.53%	2.07%	97.40%	6427

Table 3.4: Migration status by sectors

Migration type	Male			Total	Female			Total
	Native	Intra-state migrants	Inter-state migrants		Native	Intra-state migrants	Inter-state migrants	
<b>Occupation</b>								
Legis	0.24%	1.95%	97.8%	7134	0.81%	2.03%	97.15%	246
Prof	0.14%	2.34%	97.52%	5119	0.38%	1.50%	98.12%	266
Tech	0.09%	1.26%	98.65%	4376	0.23%	1.64%	98.13%	428
Clk	0.26%	1.71%	98.03%	3040	0%	1.43%	98.57%	140
Svs	0.13%	1.30%	98.56%	9755	0.31%	1.57%	98.12%	638
Agr	0.22%	1.04%	98.74%	18443	0.25%	0.83%	98.92%	1202
Crft	0.15%	1.15%	98.70%	10739	0%	0.83%	99.17%	484
Plnt	0.25%	1.19%	98.56%	4781	0%	3.85%	96.15%	52
Elemt	0.28%	1.77%	97.95%	16444	0.22%	1.84%	97.93%	1790
Niu	0.42%	2.27%	97.30%	9413	0.53%	2.06%	97.42%	6467

Legis: legislators, senior officials and manager

Prof: professionals

Tech: technicians and associate professionals

Clk: clerks

Svs: service workers and shop and market sales

Agr: skilled agricultural and fishery worker

Crft: crafts and related trades workers

Plnt: plant and machine operators and assembly operators

Elemt: elementary occupations

Niu: not in universe, unknown

Table 3.5: Labor force participation

Variables	2SLS		IV Probit			Recursive			
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.			
InterMig	0.136224	0.277698	1.108258	1.475133	0.479625	*	0.261917		
Age	0.042771	***	0.001345	0.151381	***	0.006591	0.151795	***	0.00618
AgeSqrd	-0.00054	***	0.000014	-0.001913	***	0.000076	-0.00192	***	0.000065
Distance	-8.87E-06	***	3.08E-06	-0.000045	***	0.000016	-4.7E-05	***	0.000015
GDP	-2.22E-07	***	6.95E-08	-1.05E-06	***	2.97E-07	-1.05E-06	***	3.14E-07
Cons	-0.00358		0.282857	-2.4382	*	1.469176	-1.80862	***	0.297416
Size	58818			58818					
F(5, 228)	497.84								
Prob>F	0								
R-squared	0.2512								
Root MSE	0.32147								
Wald Chi2(5)				1378.83					
Log pseudo				17442.25			-24759.3		
<b>Wald test</b>									
Chi2(1)				0.4			2.39501		
Prob>Chi2				0.5268			0.1217		
rho							-0.13043		0.08332

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 3.6: Wage-employment

Variables	2SLS		IV Probit			Recursive			
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.			
InterMig	0.84503	0.555328	2.342754	*	1.343166	-0.90377		4.001531	
Age	0.014542	***	0.001512	0.080361	***	0.005259	0.081423	***	0.007499
AgeSqrd	-0.000246	***	1.38E-05	-0.001146	***	6.84E-05	-0.00118	***	8.05E-05
Distance	0.000016	***	5.40E-06	4.27E-05	***	1.41E-05	3.47E-05	*	1.65E-05
GDP	7.76E-07	***	2.83E-07	2.05E-06	***	7.51E-07	2.17E-06	***	7.16E-07
Cons	-0.559003		0.566115	-3.730559	***	1.299886	-0.50701		4.10508
Size	58688			58688					
F(5, 228)	450.2								
Prob>F	0								
R-squared	0.0433								
Root MSE	0.48012								
Wald Chi2(5)				1834.67					
Log pseudo				1512.403			-40710		
<b>Wald test</b>									
Chi2(1)				3.12			0.040452		
Prob>Chi2				0.0775			0.8406		
rho							0.367981		1.659684

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 3.7: Day laborer

Variables	2SLS		IV Probit		Recursive	
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.
InterMig	0.793401	0.536169	2.600216 *	1.541079	1.199887 ***	0.287051
Age	0.003272 ***	0.001123	0.035967 ***	0.00491	0.035367 ***	0.004835
AgeSqrd	-0.000081 ***	1.05E-05	-0.000573 ***	6.13E-05	-0.00058 ***	5.16E-05
Distance	-3.E-06	4.98E-06	-1.04E-05	1.69E-05	-1.4E-05	1.62E-05
GDP	-1.03E-06 ***	6.49E-08	-5.09E-06 ***	5.55E-07	-5.33E-06 ***	4.59E-07
Cons	-5.15E-01	0.535293	-3.599216 ***	1.463813	-2.22478 ***	0.268256
Size	58688		58688			
F(5, 228)	261.33					
Prob>F	0					
R-squared						
Root MSE	0.40878					
Wald Chi2(5)			1234.42			
Log pseudo			9651.41		-32540.2	
<b>Wald test</b>						
Chi2(1)			3.39		12.5232	
Prob>Chi2			0.0657		0.0004	
rho					-0.6808	0.125926

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 3.8: Unpaid family workers

Variables	2SLS		IV Probit		Recursive	
	Coef.	Robst S.E.	Coef.	Robst S.E.	Coef.	Robst S.E.
InterMig	-0.03063	0.019952	-4.203683 **	1.99732	-2.17466 ***	0.823153
Age	-0.000036	8.43E-05	0.069501 ***	0.026215	0.088208 ***	0.026543
AgeSqrd	0.000002 **	9.76E-07	-0.000454 **	0.000204	-0.00057 ***	0.000218
Distance	-1.E-07	2.73E-07	-1.54E-05	3.35E-05	-5.14E-06	3.65E-05
GDP	-1.80E-09	5.63E-09	-2.34E-07	8.58E-07	-3.55E-07	9.91E-07
Cons	2.97E-02	0.020597	-0.520509	2.950714	-3.62887 ***	1.084658
Size	58688		58688			
F(5, 228)	10.94					
Prob>F	0					
R-squared						
Root MSE	0.04786					
Wald Chi2(5)			89.84			
Log pseudo			36867.98		-5340.14	
<b>Wald test</b>						
Chi2(1)			3.84		9.57541	
Prob>Chi2			0.0492		0.002	
rho					0.75075	0.137448

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.



Table 3.9: IV relevance

Variables	InterMig								
	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.	Robst S.E.	
InterRto	0.144816	***	0.039220	0.144816	***	0.005925	6.576012	***	0.202719
Age	-0.000749	***	0.000266	-0.000749	***	0.000220	-0.031642	***	0.007355
AgeSqrd	0.000005	**	0.000002	0.000005	**	0.000001	0.000236	***	0.000092
Distance	-0.000002		0.000003	-0.000002	***	0.000001	-0.000041		0.000001
GDP	1.35E-07	**	6.38E-08	1.35E-07	***	1.53E-08	5.04E-06		2.4E-07
Cons	0.940968	***	0.018688	0.940968	***	0.006008	0.270804		1.92E-07
Size	58818			58818					
F(5, 58812)	3959.27								
Prob>F	0								
R-squared	0.251								
Root MSE	0.25146								
Wald Chi2(5)				10232.72					
Log pseudo				17443.46			-24758.13		
<b>Wald test</b>									
Chi2(1)				3.02			2.39141		
Prob>Chi2				0.082			0.122		
rho							-0.130656		0.083525

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 3.10: Summary of Signs of Robustness Checks

Labor Force Participation						
Category	2SLS		IV Probit		Recursive	
	Significance	Sign	Significance	Sign	Significance	Sign
Male		-		-		+
Female	**	+	***	+	***	+
Male < 30		-	***	-	***	-
Female < 30		+	**	+		+
Wage-employment						
Category	2SLS		IV Probit		Recursive	
	Significance	Sign	Significance	Sign	Significance	Sign
Male		+		+		-
Female	***	+	***	+	***	+
Male < 30		-		-		-
Female < 30		+	**	+		+
Emplorer						
Category	2SLS		IV Probit		Recursive	
	Significance	Sign	Significance	Sign	Significance	Sign
Male		-		-	**	-
Female		-		-		+
Male < 30		-		-		-
Female < 30		+				
Day Laborer						
Category	2SLS		IV Probit		Recursive	
	Significance	Sign	Significance	Sign	Significance	Sign
Male		+		+		
Female	**	+	***	+	***	+
Male < 30		-		-		-
Female < 30		-		-	***	+
Unpaid Family Worker						
Category	2SLS		IV Probit		Recursive	
	Significance	Sign	Significance	Sign	Significance	Sign
Male		-				
Female		-		-	**	-
Male < 30		-			***	+
Female < 30						-

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

A blank row means there is no regression result to show.

Table 3.11: Female Labor Force Participation

Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
InterMig	1.607080	**	0.663418	3.971718	***	1.106505	1.003472	***	0.406970
Age	0.040718	***	0.002699	0.126666	***	0.016841	0.152548	***	0.009186
AgeSqrd	-0.000473	***	0.000026	-0.001459	***	0.000194	-0.001768	***	0.000096
Distance	-0.000017		0.000013	-0.000039		0.000031	-0.000048	*	0.000029
GDP	-2.78E-07		2.41E-07	-7.66E-07		6.14E-07	-7.61E-07		6.07E-07
Cons	-1.868528	***	0.66802	-6.39433	***	0.877016	-3.98192	***	0.450036
Size	7183			7183					
F(5, 228)	93.51								
Prob>F	0								
R-squared									
Root MSE	0.52099								
Wald Chi2(5)				342.35					
Log pseudo				-904.238			-5099.31		
<b>Wald test</b>									
Chi2(1)				9.44			4.11407		
Prob>Chi2				0.0021			0.0425		
rho							-0.38564		0.170677

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 3.12: Female Wage-employment

Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
InterMig	1.632651	***	0.646620	4.398125	***	1.084459	1.205624	***	0.340605
Age	0.022446	***	0.002222	0.097838	***	0.014644	0.125968	***	0.009587
AgeSqrd	-0.000274	***	0.000021	-0.001168	***	0.000174	-0.001515	***	0.000103
Distance	0.000002		0.000021	0.000003		0.000029	0.000003		0.000029
GDP	4.12E-07		3.24E-07	1.05E-06		8.36E-07	1.54E-06		8.83E-07
Cons	1.740239	***	0.653044	-6.63965	***	0.846327	-4.18024	***	0.380283
Size	7157			7157			7157		
F(5, 228)	70.27								
Prob>F	0								
R-squared									
Root MSE	0.48361								
Wald Chi2(5)				361.79					
Log pseudo				-183.217			-4356.21		
<b>Wald test</b>									
Chi2(1)				15.07			8.24299		
Prob>Chi2				0.0001			0.0041		
rho							-0.6059		0.154838

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

Table 3.13: Female Day Laborer

Variables	2SLS			IV Probit			Recursive		
	Coef.		Robst S.E.	Coef.		Robst S.E.	Coef.		Robst S.E.
InterMig	1.569780	**	0.638196	4.754670	***	1.090644	1.425616	***	0.135357
Age	0.014910	***	0.001824	0.075750	***	0.014565	0.100088	***	0.009809
AgeSqrd	-0.000163	***	0.000018	-0.000834	***	0.000161	-0.001106	***	0.000102
Distance	-0.000010		0.000010	-0.000030		0.000031	-0.000043		0.000034
GDP	-6.61E-07	***	1.61E-07	-3.23E-06	***	1.10E-06	-4.41E-06	***	1.18E-06
Cons	-1.672421	***	0.638306	-6.85633	***	0.80772	-4.36413	***	0.238171
Size	7151			7157			7157		
F(5, 228)	26.57								
Prob>F	0								
R-squared									
Root MSE	0.41897								
Wald Chi2(5)				237.64					
Log pseudo				730.7285			-3429.07		
<b>Wald test</b>									
Chi2(1)				17.36			29.7464		
Prob>Chi2				0			0		
rho							-0.8884		0.054649

\*: significant at 10%. \*\*:significant at 5%. \*\*\*:significant at 1%.

## BIBLIOGRAPHY

- [1] Agnihorti Indu, Mazumdar Indrani, and Neetha N. *Migration and Gender in India* Economic and Political Weekly, March 9, Vol. XLVIII No. 10, 2013
- [2] Allayannis, G., 1996. *Exchange rate exposure revisited*. Working Paper (DSWP-97-06), Darden Graduate School of Business Administration, University of Virginia.
- [3] Allayannis George, Ihrig Jane, and Weston James P., 2001. *Exchange rate hedging: Financial versus operational strategies*. The American Economic Review, Vol. 91, No. 2.
- [4] Allayannis George, Lel Ugur, and Miller Darius P., 2003. *Corporate governance and the hedging premium around the world*. working paper.
- [5] Allayannis George, and Ofek Eli, 2001. *Exchange rate exposure, hedging, and the use of foreign currency derivatives*. Journal of International Money and Finance 20, 273-296.
- [6] Ameer Rashid, 2010. *Determinants of corporate hedging practices in Malaysia*. International Business Research, DOI: 10.5539/ibr.v3n2p120 Source: DOAJ.
- [7] Amihud, Y., 1993. *Evidence on exchange rates and the valuation of equity shares*. In: Amihud, Y., Levich, R. (Eds.), *Exchange rates and corporate performance*. Business One, Irwin, IL.
- [8] Aretz Kevin, and Bartram Sohñke M., 2010. *Corporate hedging and shareholder value*. The Journal of Financial Research, Vol. XXXIII, No. 4, 317-371.
- [9] Arif GM, Irfan M. *Return migration and occupational change: the case of Pakistani migrants returned from the Middle East*. Pakistan Dev Rev. 1997;36(1):1-37.
- [10] Asturias Jose, Ramos Roberto, and Santana Manuel García *Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India* Journal of the European Economic Association, 17(6): 1881-1940, 2019
- [11] Bhagat Ram B. *Internal Migration in India: Are the Underclass More Mobile?* Paper presented in the 26<sup>th</sup> IUSSP General Population Conference held in Marrakech, Morocco, 27 September - 2 October, 2009

- [12] Black R, Costaldo A. *Return migration and entrepreneurship in Ghana and Côte D'Ivoire: the role of capital transfers*. Tijdschr Econ Soc Ge. 2009;100(1):44–58.
- [13] Chan Kam Wing *The Global Financial Crisis and Migrant Workers in China: 'There is No Future as a Labourer; Returning to the Village has No Meaning'* International Journal of Urban and Regional Research, Volume 34.3 September 2010 659–77
- [14] Castaldo Adriana, Deshingkar Priya, and McKay Andy *Internal Migration, Remittances, and Poverty: Evidence from Ghana and India* Working Paper, 2012
- [15] Chinese Household Income Project Survey 2013
- [16] Collins Robert, 2013. *Textual risk disclosures and investors' risk perceptions*. Review of Accounting Study 18, 1088-1122, DOI 10.1007/s11142-013-9228-9.
- [17] Demurger Sylvie, Xu Hui *Return Migrants: The Rise of New Entrepreneurs in Rural China* World Development Vol. 39, No. 10, pp. 1847–1861, 2011
- [18] Dellas Harris, and Zilberfarb Ben-Zion, 1993. *Real exchange rate volatility and international trade: A reexamination of the theory*. Southern Economic Journal, Vol. 59, No. 4, 641-647.
- [19] Djankov Simeon, Qian Yingyi, Ronald Gérard, and Zhuravskaya Ekaterina *Who Are China's Entrepreneurs?* The American Economic Review, Vol. 96, No. 2 (May, 2006), pp. 348-352
- [20] Frjters Paul, Kong Tao, and Meng Xin *Migrant Entrepreneurs and Credit Constraints Under Labour Market Discrimination* IZA Discussion Paper No. 5967, 2011
- [21] Green, W.H. *Econometric Analysis*, 6<sup>th</sup> ed. NJ: Prentice Hall, 2008
- [22] Google Map India
- [23] IPUMS International Survey India
- [24] Jorion, P., 1990. *The exchange rate exposure of U.S. multinationals*. Journal of Business 63, 331-345.

- [25] Kaur Ravinder *Across-Region Marriages: Poverty, Female Migration and the Sex Ratio* Economic and Political Weekly, Vol. 39, No. 25, pp. 2595 - 2603, 2004
- [26] Kone Zovanga L., Liu Maggie Y., Mattoo Aaditya, Ozden Caglar, and Sharma Shiddharth *Internal Borders and Migration in India* Journal of Economic Geography, 18, pp. 729 - 759, 2017
- [27] Kravet Todd, and Muslu Volkan, 1997. *Toward a positive economic theory of hedging*. American Journal of Agricultural Economics, Vol. 79, No. 2, 488-499.
- [28] Mahapatro Sandhya Rani *The Changing Pattern of Internal Migration in India: Issues and Challenges* International Conference of International Union for Study of Population, 2009
- [29] Monfardini Chiara and Radice Rosalba *Testing Exogeneity in the Bivariate Probit Model: A Monte Carlo Study* Oxford Bulletin of Economics and Statistics, 70, 2, 0305-9049, 2008
- [30] Marchant Katrin, Naude Wim, and Siegel Melissa *Migration, entrepreneurship and development: critical questions* IZA Journal of Migration (2017) 6:5
- [31] National Bureau of Statistics of China *Chinese Household Income Project*
- [32] Nguyen Hoa, and Faff Robert, 2002. *Can the use of foreign currency derivatives explain variations in foreign exchange rate exposure? Evidence from Australian companies*. Journal of Multinational Financial Management 13, 193-215.
- [33] Nydahl Stefan, 1999. *Exchange rate exposure, foreign involvement and currency hedging of firms: some Swedish evidence*. European Financial Management, Vol. 5, No. 2, 241-257.
- [34] Papaioannou Michael G., *Exchange rate risk measurement and management: Issues and approaches for firms*. South-Eastern Europe Journal of Economics 2, 129-146.
- [35] Piracha M, Vadean F. *Return migration and occupational choice: evidence from Albania*. World Dev. 2010;38(8):1141-C55.
- [36] Singh D.P. *Migration and Occupation in Mumbai: Issues and Implications* European Population Conference, 13- 16 June, Stockholm, Sweden, 2012

- [37] Ugur Lel, 2012. *Currency hedging and corporate governance: A cross-country analysis*. *Journal of Corporate Finance* 18, 221-237.



VITA

YIMING LIU

2011	B.A., Business English(International Trade) Wuyi University Jiangmen, China
2013	M.S., Finance Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

Eastern Economic Association Annual Conference 2020

Southern Economic Association Annual Conference 2020