August 2016

Essays on Applied Microeconomics

Jin Zhou
The University of Western Ontario

Supervisor
Salvador Navarro
The University of Western Ontario

Joint Supervisor
Nirav Mehta
The University of Western Ontario

Graduate Program in Economics

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

Abstract

My dissertation aims to contribute the migration and education literatures, my thesis contains two main topics and three papers. In the first two papers, I study the job search process of rural migrants in urban labor markets when social networks present. Then, I analyse how social networks affect both migration decisions and individuals' labor market outcomes. In the last paper, I examine the determinants of an individual's college education decision, with particular attention to the uncertainty faced by individuals.

The first paper examines the job search process of rural migrants, who have the option of returning home. This paper focuses on analysing the effect of social networks on their labor market outcomes. I build a dynamic structural model of job search for migrants. I estimate the model using "Rural Urban Migration in China" dataset. The estimation results show that rural migrants with social networks receive more job offers and that migrants with social networks also have higher wages on average.

The second paper analyses the effect of social networks on both migration decisions and individuals' labor market outcomes. I develop and estimate a dynamic model of return and repeated migration, social network investment decisions and labor market transitions. The model distinguishes two channels through which social networks may affect migration decisions: (1) a direct effect on migration costs, and (2) an indirect effect on labor market outcomes through the job arrival rate. I estimate the model using panel data from the Chinese Household Income Project (2007-2009). The estimation results show that social networks increase arrival rates, and decrease migration costs. I also show that policies that directly lower migration costs may be more cost effective at increasing rural-urban migration in China than those focused on unemployment benefits.

In the third paper, we use economic theory and estimates of a semiparametrically identified structural model to analyze the role played by uncertainty and its interaction with credit constraints and preferences in explaining education choices. We develop a methodology that distinguishes information unknown to the econometrician but forecastable by the agent and information unknown to both at each stage of the life cycle. Using microdata on earnings, we estimate a model of college choice, labor supply and consumption, under uncertainty with repayment constraints. We find that 52% and 56% of the variances of college and high school log-wages respectively are predictable by the agent at the time schooling choices are made. When people are allowed to smooth consumption, college increases to nearly 58%.

Keywords: Internal Migration, Social Networks, Search, Education Choice, Uncertainty, Borrowing Constraint

Co-Authorship Statement

This thesis contains material co-authored with Salvador Navarro. All the authors are equally responsible for the work appears in Chapter 3 of this thesis.

Acknowlegements

I would like to express my deepest gratitude to my thesis supervisors, Professor Salvador Navarro and Nirav Mehta, for their invaluable guidance and tremendous support throughout my graduate program. Discussion with them helps me to learn from their economic insights and their encouragement has been a constant motivation for me to carry on the work in this thesis.

I would also like to express the deepest appreciation to my thesis committee members: Professor Audra Bowlus and Professor Terry Sicular for the many hours they spent developing my research and pushing me to do better. They also give me many great suggestions about teaching, which will benefit me for my future academic career.

I would also like to acknowledge the faculty members in the Department of Economics for their time and active participation at seminars and presentations. I would like to especially thank Chris Robinson, Lance Lochner, Todd Stinebrickner and David Rivers for their helpful comments. I also thank seminar participants in the University of Western Ontario and conference participants at 2015 4^{th} Society of Labor Economists World Meeting, 2015 Annual Conference of Canadian Economics Association and 2014 UM-MSU-UWO Labor Day Conference. This research would not have been possible without financial support from the Social Sciences and Humanities Research Council of Canada.

Dedicated to my parents

Contents

List of Figures

List of Tables

List of Appendices

Chapter 1

Do Networks Matter? Job Search with Return Migration in China

1.1 Introduction

Nowadays, one of the largest internal migration in the human history is happening in China. From the investigation conducted by National Bureau of Statistics of China (NBSC), more than 132 million people who were born in rural areas had migrated in 2006 to urban cities, which is triple the number of migrating Europeans to the United States since industrialization. Many scholars in China point out that this number has been underrated. For instance, Cai (2010) estimated the number of migrants to be 188 million. The recent report by National Bureau of Statistics of China (2012) mentioned that the number of rural migrants was above 253 million.

There are two main features of rural-out migration in China: networks have a significant correlation with the job search process, i.e., migrants with networks have around 40% shorter unemployment duration than those without networks. Another feature of Rural-Urban migration in China is that it is temporary. Murphy (1999) shows that almost one third of migrants returned to their original location in China. Deng *et al* (2009) mentions only 8.13% of rural migrants (State Council Research Bureau 2006) declared that they planned a long term stay at the destination city.

Rural migrants have become an enormous part of the labor market in urban cities. Identifying the factors which affect rural migrants' transitions into the urban labor market is necessary to understand the phenomenon of a rural migrants' job search process. Also it would help to analyze related policy issues, such as the crime control policy and new social welfare programs for rural migrants.

The paper develops a dynamic job search model. To capture the two features of rural-out migration in China, first, this paper considers the effect of networks on both job arrival rates and migrants' wages; second, to model the behaviour of return migration, I allow individuals to have the option of return migration in this paper. This model allows unemployed migrants to have the permanent possibility to quit searching for jobs and return to their home town. This environment extends the job search framework of Van den Berg (1990) which does not assume unemployed workers have the option to leave the labor market.

It is important to add the option of return migration when we consider the job search of migrants. When we study the transition from unemployment to employment, it may cause a biased estimation if we fail to take the option of return migration into account. First, unemployed migrants with lower expected returns to job search may quit job searching earlier than unemployed migrants with higher expected returns of job search. Second, a higher option value of return migration may urge migrants to go back more quickly than others with lower option values. The transition from unemployment to return migration is not independent on the probabilities of the transition from unemployment to employment, which cannot be treated as an exogenous right-censoring of duration of unemployment.

In this paper, I also will examine the role of networks on job arrival rates and wage distributions. From the data[1], it is found that migrants with networks are more likely to find jobs but they have lower wages. Networks help migrants increase the probability of receiving job offers. More information may also help migrants to have a better match with firms. Also migrants with networks get higher wages. This paper tries to analyze the effects of networks on job arrival rates and wage distributions. It helps us to know more about how networks affect the labor market outcomes, especially in developing countries.

The contribution of this paper is to provide a structural framework of the analysis of the flow of migrants among different labor market states (employment, unemployment, and return migration). It is also the first time to estimate the dynamic job search process of rural migrants in China. This study could help the government to build an efficient social welfare program for rural migrants. The empirical results provide the structural estimates of the effects of networks on job search and the empirical results also help us to understand how the option of returning home affects rural migrants' job search behaviors. Second, the data offers the best statistics regarding migration in China. It is the nationalized data source, which includes 9 provinces covering more than 50 percent of rural migrants in China. The national data can provide more correct information of rural migrants' performances in urban cities.

Section I is the introduction of the article. Section II gives a short literature review of related topics. I will introduce the model and the interpretation of the dynamic process in section III. Section IV gives the information of data source and the statistics of key variables. The estimation results are shown in the section V.

[1] I use first two round of Rural and Urban Migration in China to estimate the model in this paper.

Section VI concludes.

1.2 Literature review

1.2.1 Search model

The seminal work of the search model is Kiefer and Neuman (1979), which was the first to empirically investigate implication of the search model with constant reservation wage , using reduced-form equations. Flinn and Heckman (1982) was the first to structurally estimate the model. In the model, I consider the return migration as an option of rural migrants when they are searching for jobs in the urban market. Hence, the model I consider is related to the topic of job search with nonparticipation.

The empirical literature which considers nonparticipation is based on the approach of treating transitions from unemployment to nonparticipation as stochastic occurrences. Banerjee (1984) analyzed a general framework which makes it possible to examine an individual's inter-temporal discrete choice decisions. In that work, an individual will choose to occupy one of the three labor market states at each moment faced and their expectations about future changes in the environment. Mortensen and Neuman (1984) compare Banerjee (1984) and Flinn and Heckman (1982), and then estimate the model of Banerjee (1984).

The recent work of Van den Berg (1990) and Frijters and van der Klaauw (2006) consider estimating the structural search model with nonparticipation. Van den Berg (1990) estimates a stationary search model with nonparticipation but in his model he assumes the income flow after becoming nonparticipation is close to the constant benefit level. This assumption helps him fix the transition rate from an unemployment to nonparticipation state. In this paper, this assumption will be relaxed (i.e. the value of nonparticipation is not a function of benefit (cost) level). Frijters and van der Klaauw (2006) consider job search with nonparticipation model within the environment of non-stationarity. The path of reservation wages is dependent on the value of nonparticipation, which is unobserved. They have to use a simulated likelihood method to estimate the model.

1.2.2 Return migration

Most economic research treats migration as permanent decision (i.e., Chiswick (1978) and Borjas (1999)). However, levels of return migration have been quite high. For example, Jasso and Rosenzweig (1982) point out that the 1971 cohort of immigrants to the US, almost fifty percent returned to their home country by 1979. Research which studies return migration is sparse. Galor and Stark (1990) examine migrants with the probability of returning prefer to save more money than comparable native-born people. An important contribution to the theoretical explanations of return migration is the work of Borjas and Bratsberg (1996). They argued that return migration may have been planned as part of an optimal life-cycle residential location sequence. Return migration also occurs because immigrants based their initial migration decision on erroneous information about opportunities in the host country.

For rural-urban migration in China, the level of return migration is noticeable. It is necessary to consider return migration when we analyze the behavior of migrants. In the 1990's, as the outflow of rural migrants increased, return migration also became noticeable. About one-third of all migrants are estimated to return to their original places (Murphy (1999)).

Hare (1999)'s paper is the first to consider return migration in China. She used the MPH model with weibull distribution to estimate the duration of migrants in cities. She found that more land allocations in rural areas would decrease the duration in cities. More female workers in the household could increase the spell in cities. Zhao (2002) found married migrants had a higher probability of return as well migrants with a higher education level also had a higher likelihood to return. Hare (1999) used the reduced form duration model and Zhao (2002) applied the binomial logit model to estimate the relationship between key variables and the return migration decision. Deng *et al* (2009) mentions only 8.13% rural migrants (State Council Research Bureau 2006) declared that they planned a long term stay at the destination city and they examine the land rights insecurity is the main reason of temporary migration in China.

But all of them ignore the search process of migrants in cities. The decision of return migration should be made after comparing the value of unemployment with the value of employment. We did not have information about migrants' performances in destination labor markets in those two papers. It could be some reasons which cause that the returned migrants have a higher value of returning or some reasons which cause those returned migrants have lower job arrival rates, higher search costs or lower wages. Hence, we need a model to give answers to the above questions and to understand migrants' labor market performance in the urban city.

1.2.3 Social Network and Migration

The role of information on the destination labor market may be crucial for success. The information problem for migrants may be bigger both in distance and in culture. The opportunity cost of remaining in the host is lower for close by. For example, Borjas and Bratsberg (1996) find out immigrants to the US tend to return to rich and to countries close to the US. Ethnicity is also important if immigrants of a certain ethnic group systematically perceive a lower return than expected. Not only for migration across countries, social networks but also play the most important role for internal migration. Immigrants can get more useful information from their own social networks. Although the role of social networks on migration has been widely studied (i.e Banerjee (1983),Banerjee (1984); Munshi (2003)), how networks influence unemployment, especially considering return migrations are very scarce.

Recent studies about the Rural-Urban migration in China mainly focus on network effects on migration decision-making (i.e Hare and Zhao (2000); Meng (2000); Zhao (2003)). Zhang and Zhao (2011) examined the relationship between social-family networks and self-employment and they find social-family networks would help migrants to be self-employed in the destination cities. However, little is known about the role of migration networks in determining labor market outcomes like wages, job arrival rates, spells of job search. If we ignore networks' effect, it is possible that we may obtain a biased understanding of migrants' behaviors in the labor market.

1.3 Model

1.3.1 Outline of the Model and Likelihood

In this section, I present the structural framework to model the transition from unemployment to employment and to return migration. This model is based on the work of the continuous time job search theory (Mortensen (1986)) and extends the model of Van den Berg (1990) which estimate the job search model considering the probability of nonparticipation. Here, I allow the value of return migration to be a random variable, which does not need to be constant as it is assumed in Van den Berg (1990). But in our framework, job arrival rates, wage distributions and the distribution of values of return migration still do not change over the elapsed duration of unemployment. This means our model is in the stationary environment. After providing the outline of the model, the identification is discussed briefly.

In this model, unemployed migrants search sequentially for jobs until a suitable wage offer has been found or the value of return migration is large enough. Job offers arrive randomly in time at rate λ. Wage is drawn randomly from a wage distribution $F(w)$, which is assumed to be log-normal distribution; the value of return migration is randomly drawn from the distribution $G(\eta)$. During the unemployment time, migrants have to pay the cost of search c (i.e., the transportation fee and the rent of apartment). In China, there is no law to require the local government to pay unemployment insurance or benefits to rural migrants. In our data, only less than 1% of rural migrants do receive unemployment benefits, so in this paper, I only assume unemployment migrants have to pay the search cost and cannot get any unemployment benefits.

The variables λ, c and w are measured per unit period (day). Each migrant aims to maximize their expected discounted lifetime income. I assume once a job is accepted it is kept forever at the same wage and if they make decisions to return, they cannot search for jobs in the urban labor market any more. This assumption is not too strong since there are quiet few on-the-job search behaviour for rural migrants.

I consider the model in the stationary environment, which means λ, c, $F(w)$, $G(\eta)$ are independent of the unemployment duration. This assumption is not very realistic. Because job arrival rates may decrease during the unemployment duration. Wage distribution and the distribution of values of returning would change due to business cycle effects. In the data, I do not have the detailed information including the reservation wage, social benefits across different unemployment durations. This means that I cannot build up the non-stationary model. But the duration of job search for rural migrants in China is quite short compared with job search duration in European countries. The mean of unemployment duration of rural migrants is around 28 days. In such a short time, the wage distribution and the value of returning distribution can be treated as constant. This may not hurt the understanding of the job search process of rural migrants very much.

In this paper, I use the optimal stopping search model to explain the job search process. The search model also will help us to identify job arrival rates and distributions of wages and values of return migration. I also can examine the effect of social networks on labor market outcomes (i.e., job arrival rates, and wages).

In the model, an unemployed migrant who receives a job offer with probability λ and has to decide whether to accept it, to return migration or to continue searching in the hope of obtaining a better offer in the future. If the unemployed migrant does not receive a job offer, he has to make a decision whether to continue search or return home.

Let V_u denote the expected present value of unemployment, V_e denote the expected present value of employment and V_r denote the expected present value of return migration. The Bellman equation for V_u satisfies:

$$V_u = c + \frac{\lambda}{1+r} E_{w,\eta} \max\{V_e, V_u, V_r\} + \frac{1-\lambda}{1+r} E_{\eta} \max\{V_u, V_r\} \tag{1.1}$$

In continuous time of search set up, the Bellman equation 1.1 can be rewritten as:

$$V_u = c + \frac{\lambda}{1+r} E_{w,\eta} \max\{\frac{w}{r}, \frac{\eta}{r}, V_u\} + \frac{1-\lambda}{1+r} E_{\eta} \max\{V_u, \frac{\eta}{r}\}$$

If the unemployed migrant receives a job offer, he compares three options: taking the job offer, continuing to search for a better one, and the option of return migration. If the migrant does not receive a job offer, he compares the value of continued unemployment with the value of return migration. From the Bellman equation, it is easy to see if the expected value of unemployment exceeds the present value of return migration, migrants will continue to search for a job. Once the wage offer is high enough (i.e., the value of employment is larger than the value of unemployment and the value of return migration), migrants will accept the job offer.

The difference of this model as compared to others is that the optimal strategy of an unemployed individual can be characterized by the value of unemployment not by the fixed reservation wage. Because at each period, the value of return migration is randomly drawn from the distribution of the value of return, the reservation wage is not a fixed number over the unemployment duration. That means at each period, the reservation wage will reflect the information of the value of return migration.(i.e., a higher draw of the value of return will cause a higher value of reservation wage at that period.) In this sense, this model is different from other stationary search models. The migrants' decisions are based on the value of unemployment and two random variables: values of wages and return migration.

The following theorem shows that the solution of the Bellman equation (3.1) has an unique solution when the search cost exists (i.e., $c < 0$).

Theorem 1.3.1 *The function $\phi(V_u)$ is monotonically decreasing. Furthermore, if $c < 0$, the function $\phi(V_u) = 0$ has an unique root.*
where, $\phi(V_u) = c + \frac{\lambda}{1+r} E_{w,\eta} \max\{V_e, V_u, V_n\} + \frac{1-\lambda}{1+r} E_\eta \max\{V_u, V_n\} - V_u$[2]

The transition rate from unemployment into employment θ_{ue} can be written as the product of the job arrival rate and the probability of accepting a job offer:

$$\theta_{ue} = \lambda \Pr(V_e > max\{V_u, V_r\}) \tag{1.2}$$

Similarly, the transition rate from unemployment into return migration θ_{ur} can be described as the sum of two terms. The first one is the product of the job arrival rate and the probability of returning. The second term is the product of the probability without job offers and the probability of the value of returning is bigger than the value of continue search:

$$\theta_{ur} = \lambda \Pr(V_r > max\{V_e, V_u\}) + (1 - \lambda) \Pr(V_r > V_u) \tag{1.3}$$

Compared our model with the model of Van den Berg (1990), in Van den Berg(1990), the data of nonparticipation is not available so he just assumes the transition from unemployment into non-participation can be characterized by a Poisson process with a parameterized transition rate ζ. In the model, based on the Heckman and Singer (1984), I can identify the distribution of the value of return. While in Van den Berg (1990), ζ is just a exogenous parameter. Also in Van den Berg (1990), to make sure to get a fixed transition rate from unemployment to nonparticipation, he assumed the value of nonparticipation equals to the benefits of unemployment. While in this model, I do not need this assumption. Hence, this model has a weaker assumption.

[2]The proof is in appendix A.

Likelihood Function

The likelihood function is associated with the hazard function. As we all know, the hazard function $h_i(t)$ is defined as:

$$\begin{aligned} h_i(t) &= \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t | T \ge t)}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{P(t \le< t + \Delta t \cap T \ge t)}{P(T \ge t)\Delta t} \\ &= \frac{1}{P(T \ge t)} \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t)}{\Delta t} \\ &= \frac{f_i(t)}{S_i(t)} \end{aligned}$$

where $f_i(t)$ is density function of the events (i.e. accepting a job or return migration): $f_i(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t+\Delta t)}{\Delta t}$ and $S_i(t)$ is survival function: $S_i(t) = e^{-\int_0^t h_i(u)du}$.

Hence the density function of duration t can be written as:

$$f_i(t) = h_i(t) S_i(t)$$

Accepting a job offer and returning migration are two competing risks, so the hazard function for each individual should be the sum of those two transition rates. $h_i(t) = \theta_{i,ue} + \theta_{i,ur}$. Then the density function of employment is :

$$\begin{aligned} f_{i,ue}(t) &= f_i(t) \frac{\theta_{i,ue}(t)}{h_i(t)} \\ &= h_i(t) e^{-\int_0^t h_i(u)du} \frac{\theta_{i,ue}(t)}{h_i(t)} \\ &= \theta_{i,ue}(t) e^{-\int_0^t (\theta_{i,ue}(t) + \theta_{i,ur}(t))du} \end{aligned}$$

Due to the property of stationarity, the transition rate is independent with the unemployment duration. So the density of employment can be written as:

$$f_{i,ue}(t) = \theta_{i,ue} e^{-(\theta_{i,ue} + \theta_{i,ur})t} \tag{1.4}$$

Similarly, the density of return migration is:

$$f_{i,ur}(t) = \theta_{i,ur} e^{-(\theta_{i,ue} + \theta_{i,ur})t} \tag{1.5}$$

The likelihood function of this model[3] is

$$L = \begin{cases} \text{employed} & \theta_{i,ue} e^{-(\theta_{i,ue} + \theta_{i,ur})t} \\ \text{return} & \theta_{i,ur} e^{-(\theta_{i,ue} + \theta_{i,ur})t} \\ \text{unemployed} & e^{-(\theta_{i,ue} + \theta_{i,ur})t} \end{cases} \tag{1.6}$$

The density function $f(t)$ and transition rates θ are the functions of the value of unemployment V_u. Once the value of unemployment for each individual is determined, the value of likelihood can be calculated.

[3]The details of likelihood in the Appendix B

1.3.2 Some Remarks on the Identification

In this section, the identification of the structural elements of the model is briefly discussed. These structural parameters are the wage offer distribution $F(w)$ with the variance σ_w^2, the job arrival rate λ, the distribution of the value of return migration, $G(\eta)$.

From the accepted wages after an unemployment spell, I can identify the wage distribution which satisfies $F(w|V_e > \max\{V_u, V_r\})$. The variance σ_w^2 is determined based on the assumption of the shape of wage distribution. In this paper, I assume $F(w)$ follows the log-normal distribution.

The transition rates θ_{ue}, θ_{ur} can be identified from the observed unemployment durations. Because I assume the shape of the wage distribution is log-normal, job arrival rates λ can be identified separately from the wage distribution.

The distribution of return migration is identified from the duration of the transition from unemployment state to return migration.

Finally the discount rate r cannot be identified in this model. I assume $r = 0.000083$ in this paper to target the annual interest rate is 3%.

1.3.3 Parametrization

This model can be estimated in continuous time if we know the value of unemployment. The theorem 1.3.1 shows the solution of the nonlinear equation is unique, which means that we can solve Bellman equation numerically and then estimate the model.

The job arrival rate is parameterized by:

$$\lambda = \frac{e^{X'_\lambda \beta_\lambda}}{1 + e^{X'_\lambda \beta_\lambda}}$$

The wage distribution comes from the $F(x_w, \epsilon)$ and the variance σ_w^2 of the wage distribution. Here I assume the wage distribution is log-normal distribution with mean $x'_w \beta_w$

$$\ln(w) = x'_w \beta_w + \epsilon$$

The search cost function is specified as $-e^{x'_c \beta_c}$. Based on the spirit of Heckman and Singer(1984), I treat the distribution of the value of return migration as a discrete distribution (η_1, η_2, η_3) with zero mean. Hence, the whole parameter space is $\beta_\lambda, \beta_c, \beta_w, \sigma_w^2, \eta_1, \eta_2, \eta_3$ and the probability p_1, p_2, p_3.

The estimation procedure is two steps. Firstly, when given the initial guess of parameters, I solve the inner loop which is the non-linear Bellman equation for each individual, to calculate each migrant's value of unemployment. Secondly, with the value of unemployment for each individual in hand, I can calculate the likelihood for each individual. The outer loop is estimated by the method of maximizing likelihood estimation.

1.4 The Data and Some Empirical Facts

This study uses the first two waves of panel survey data which is Rural-Urban Migration in China and Indonesia (RUMiCI). This database is planned to be a five-year panel survey in China and Indonesia with the goal of studying issues such as the effect of rural-urban migration on income mobility and poverty alleviation, the state of education and health of children in migration families.

The first wave of the survey was conducted in 2008. In China, three representative samples of households were surveyed, including 8000 migration households in 15 cities[4], a sample of 8000 rural households in 9 provinces and a sample of 5000 urban households in 9 provinces. In this paper, the empirical analysis uses the information from the migration and rural samples. The 15 cities cover most of the migration destinations in China. Eight of these cities are in coastal regions (Dongguan, Guangzhou, Hangzhou, Nanjing, Ningbo, Shanghai, Shenzhen and Wuxi); five of them are in central inland region (Bengbu, Hefei, Luoyang, Wuhan, and Zhengzhou); and two of them are in the west (Chengdu and Chongqing). Figure 1.1 gives the map of the 15 cities. These 15 cities are included in the same 9 provinces as rural and urban samples. In 2000 to 2005, more than 68.38% of migrants moved into those 9 provinces and while 52.10% of migrants moved out of those 9 provinces (NBS 2002, 2007). Table 1.1 shows the detailed information of the flow of migrants in China from 1990 to 2005. Figure 1.2 is generated based on the numbers in Table 1.1. We can find the dark green areas are the locations where rural people migrate out and the dark brown areas are the places which are the destinations of rural migrants. The data used by this paper cover most places where migrants move out and move in.

The migration survey includes detailed information about the respondents' personal characteristics, educational background, employment situation, health status, children's education, social and family relationships, major life events, income and expenditure, housing and living conditions. In terms of basic information of a household member, we know the person's age, gender, education level, number of children, current address and the original home town address, the time of first

[4]The 15 cities are Bengbu, Chengdu, Chongqing, Dongguan, Guangzhou, Hangzhou, Hefei, Luoyang, Nanjing, Ningbo, Shanghai, Shenzhen and Wuhan, Wuxi, and Zhengzhou

Figure 1.1: Sample Cities in China

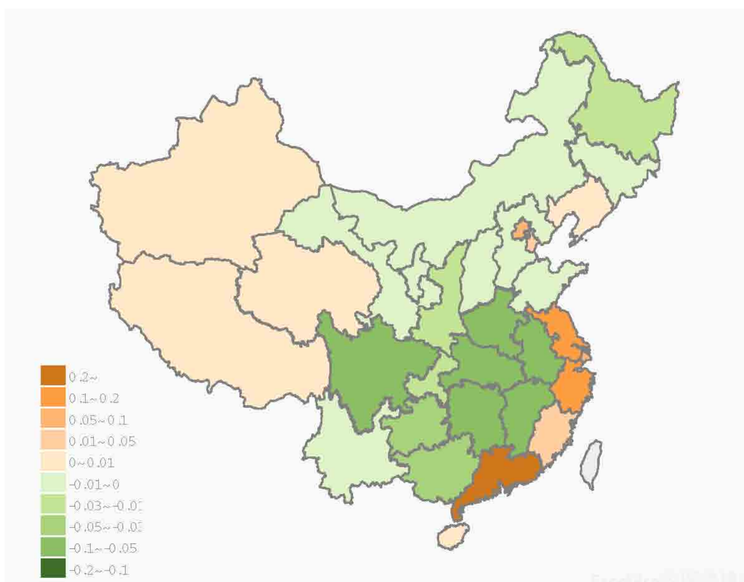


Figure 1.2: Migration flows in China From 2000 to 2005

migration, the search duration of first job, the wage of first job, the duration of staying in the city, monthly rent of house, monthly payment for transportation. Also, I know who provided the information before first time migration and we know who the migrant's most important contacts are and whether they live in the same city, and whether the migrant's parents and sibling also live in the same city.

Table 1.2 gives the key variables of this analysis. In this paper, I restrict the migrants who are older than 15 and younger than 65. The health variable is a self-reported variable which shows whether the migrant thinks he (she) is healthier than those within his(her) same cohort.

This paper considers return migration as one option choice for all migrants and the rural survey provides information to identify people who returns from migration destinations. Because the rural survey questionnaire only has information about current migration. This forces us only to consider people who first migrated in 2007 and in 2008.

In this paper return migrants are rural migrants who returned their home and then stay in their counties for at least three months. The definition of migration is the destination out of their own counties. Networks are defined in two ways: 1) migrants have relatives or friends living in the destination cities; 2) migrants who are provided the job information in destination cities. There would exist selections based on social networks. This chapter mainly focuses on the correlation between social networks and labor market outcomes. The next chapter will correct the selection of social networks.

The information of transportation and rent can only be found in the migrate data source. I calculate both the means of transportation and housing rent for each province and use those means as the proxy numbers for the samples from rural data. The dummy variables of transportation and rent are the indicators if the values of transportation or rent are proxy by the mean.

Almost half of migration households are missing (3912 households) in the first wave so 3912 new households are now included in the second wave. Because both first and second wave questionnaires ask key information in this empirical analysis, the missing data does not impact this research very much. In this paper, I only focus on first time migrants. This means I only use one year information, which is cross sectional data.

Table 1.1: Interprovincial Migration in China, 1990-2005(In Thousands)

1990-1995		migration				1995-2000		migration				2000-2005		migration			
		In	Out	Net	Net(%)			In	Out	Net	Net(%)			In	Out	Net	Net(%)
1	Guangdong	1947	221	1726	16.2	1	Guangdong	11501	438	11063	34.3	1	Guangdong	11996	1715	10281	27
2	Shanghai	726	122	604	5.7	2	Shanghai	2168	163	2005	6.2	2	Zhejiang	5062	1041	4021	10.6
3	Beijing	694	117	577	5.4	3	Zhejiang	2715	970	1745	5.4	3	Shanghai	3025	375	2650	7
4	Jiangsu	969	450	519	4.9	4	Beijing	1890	174	1715	5.3	4	Jiangsu	3290	1328	1963	5.2
5	Xinjiang	566	150	416	3.9	5	Xinjiang	1142	217	925	2.9	5	Beijing	2246	330	1916	5
6	Liaoning	435	197	239	2.2	6	Fujian	1346	625	722	2.2	6	Fujian	1934	802	1132	3
7	Tianjin	223	62	161	1.5	7	Jiangsu	1908	1241	667	2.1	7	Tianjin	908	107	802	2.1
8	Shandong	527	382	145	1.4	8	Tianjin	492	104	388	1.2	8	Xinjiang	577	182	395	1
9	Fujian	344	220	125	1.2	9	Liaoning	755	380	375	1.2	9	Liaoning	674	416	257	0.7
10	Hebei	503	417	87	0.8	10	Yunnan	733	398	335	1	10	Hainan	191	158	33	0.1
11	NeiMongol	275	249	27	0.3	11	Hainan	218	130	88	0.3	11	Ningxia	74	68	7	0
12	Shanxi	158	140	18	0.2	12	Shanxi	383	334	49	0.2	12	Tibet	26	31	-6	0
13	Tibet	38	28	10	0.1	13	Ningxia	129	87	41	0.1	13	Qinghai	74	85	-12	0
14	Hainan	104	102	2	0	14	Tibet	71	35	35	0.1	14	NeiMongol	394	417	-23	-0.1
15	Ningxia	49	54	-6	-0.1	15	Shandong	904	878	26	0.1	15	Yunnan	469	601	-132	-0.3
16	Qinghai	51	77	-25	-0.2	16	Qinghai	77	123	-46	-0.1	16	Shanxi	210	345	-135	-0.4
17	Yunnan	207	242	-35	-0.3	17	Hebei	770	872	-102	-0.3	17	Shandong	924	1123	-199	-0.5
18	Zhejiang	466	514	-49	-0.5	18	NeiMongol	325	441	-116	-0.4	18	Jilin	218	532	-315	-0.8
19	Shaanxi	163	265	-101	-1	19	Jilin	254	529	-275	-0.9	19	Gansu	118	494	-376	-1
20	Hubei	271	382	-111	-1	20	Shaanxi	423	719	-296	-0.9	20	Hebei	612	990	-378	-1
21	Gansu	140	251	-112	-1	21	Gansu	204	561	-357	-1.1	21	Shaanxi	255	827	-572	-1.5
22	Jilin	150	295	-145	-1.4	22	Heilongjiang	301	940	-639	-2	22	Heilongjiang	195	1020	-825	-2.2
23	Guizhou	152	402	-250	-2.3	23	Chongqing	448	1103	-655	-2	23	Chongqing	427	1437	-1010	-2.7
24	Jiangxi	125	514	-389	-3.6	24	Guizhou	261	1232	-970	-3	24	Guizhou	531	1766	-1235	-3.2
25	Heilongjiang	224	614	-389	-3.7	25	Guangxi	287	1838	-1551	-4.8	25	Guangxi	397	2123	-1726	-4.5
26	Guangxi	120	554	-434	-4.1	26	Hubei	606	2210	-1604	-5	26	Jiangxi	499	2476	-1977	-5.2
27	Henan	270	740	-470	-4.4	27	Henan	470	2309	-1839	-5.7	27	Hubei	501	2715	-2214	-5.8
28	Hunan	215	704	-489	-4.6	28	Jiangxi	236	2681	-2445	-7.6	28	Hunan	501	3328	-2827	-7.4
29	Anhui	155	744	-589	-5.5	29	Anhui	313	2893	-2579	-8	29	Henan	280	3433	-3154	-8.3
30	Sichuan*	395	1457	-1062	-10	30	Hunan	363	3261	-2899	-9	30	Anhui	671	3836	-3165	-8.3
						31	Sichuan	590	4396	-3806	-11.8	31	Sichuan	763	3941	-3178	-8.4
		9189	9189	0				32282	32282	0				38042	38042	0	

1. Before 2000, Chongqing is a city of Sichuan province. The data of Sichuan from 1990 to 1995 includes the information of Chongqing.
2. Source from NPSSO(1997), SC and NBS (2002,2007)
3. Net%=Net migration/National total of migration×100%.

Table 1.2: Key variables

Variables:	
Gender	Male=1, indicator
Age	Age, between 15 and 65
Marriage	Marriage indicator
Minority	Race indicator
Education	Years of formal education not including training
Health	Scale variable from 1 to 5
Network	Indicator:=1 if have networks, =0 if without networks
Rent	rent payment for apartment(per month)
Transportation	transportation fee (per month)
Duration T	$T = \min\{T_1, T_2, T_3\}$
	T_1 : time for job search and find a job
	T_2 : time for job search and then return to hometown
	T_3 : right censored
Cause D	= 1: if immigrant finds a job
	= 2: if immigrant returns to hometown
	= 3: if right censored

Table 1.3 presents the summary of key variables. From Table 1.3, we can find there are more male migrants than female migrants, especially in the group of migrants without networks. The average age of migrants is 26 years old. Because compulsory education in China takes nine years, most rural people migrate out after finishing the nine year compulsory education. An interesting finding is that the monthly wage is higher for those migrants without networks than those with networks. At the same time, there are more single migrants without networks. As we expected, migrants without networks take longer to search jobs. From the labor market states, we can find migrants with networks are more likely to be employed. When I divide the full sample based on the different labor market states (employment, unemployment, and return migration), there are several distinctive findings: more female migrants are in a state of unemployment; the migrants who have returned and are unemployed are older and less educated compared with the migrants who are employed; the marriage rate is very high among the group of returned and unemployed migrants.

The most interesting finding is that return migrants have the lowest rate of networks (0.23) and unemployed migrants have the highest rate of networks (0.66). It is possible that there exists selection across different labor market states based on the networks. Hence, social networks would be an important factor to affect individuals' labor market transition. For example, networks could help migrants to have a higher job arrival rate, which would impact migrants' decision on return migration. If we ignore the variable of networks, we may underestimate the likelihood of employment for those migrants with networks. Similarly, we will get biased

Table 1.3: Sample Summary

		Network		Labor Market Status		
	Full Sample	Network=1	Network=0	Employed	Return	Unemployed
Male	0.53	0.51	0.54	0.53	0.55	0.28
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.45)
Age	26.09	26.47	25.62	25.55	32.77	34.09
	(10.35)	(10.92)	(9.60)	(9.98)	(11.68)	(14.78)
Education	9.44	9.26	9.66	9.55	8.09	7.49
	(3.02)	(2.93)	(3.11)	(2.98)	(3.13)	(3.46)
Health	1.76	1.78	1.73	1.75	1.78	2.09
	(0.71)	(0.73)	(0.69)	(0.71)	(0.68)	(0.86)
Wages	1282.80	1233.09	1348.01	1282.80		
	(792.74)	(683.69)	(912.54)	(792.74)		
Marriage	0.32	0.35	0.29	0.30	0.61	0.66
	(0.47)	(0.48)	(0.46)	(0.46)	(0.49)	(0.48)
Minority	0.02	0.02	0.02	0.02	0.01	0.02
	(0.13)	(0.13)	(0.13)	(0.13)	(0.08)	(0.15)
Disabled	0.03	0.03	0.03	0.03	0.02	0.04
	(0.16)	(0.16)	(0.16)	(0.16)	(0.15)	(0.20)
Duration	27.95	20.99	36.46	10.25	240.86	317.89
	(74.43)	(62.58)	(86.01)	(23.11)	(90.37)	(174.85)
Hours	53.55	56.56	49.88	57.62		
	(21.41)	(20.11)	(22.38)	(16.06)		
Rent	132.62	130.34	135.40	181.43	145.56	153.33
	(165.35)	(188.00)	(132.52)	(148.84)	(55.77)	(158.20)
Transportation	33.67	32.61	34.98	31.86	33.74	23.26
	(36.42)	(42.62)	(26.93)	(27.44)	(7.81)	(28.43)
Network	0.55			0.57	0.23	0.66
	(0.50)			(0.50)	(0.42)	(0.48)
Observation	3096	1703	1393	2874	175	47

1. Hours are working hours per week.
2. Health is a self-reported variable which is scaled from 1 to 5. 1 is the best and 5 is the worst.
3. Cause describes reasons to exit the current state: d=1 if immigrants find a job and d=2 if immigrants return to their hometown
4. Wage, rent, transportation are monthly. The unit of duration is day.
5. Standard deviations are reported in parentheses.

Table 1.4: Sample Summary for Male

	Male	Network		Labor Market Status		
		1	0	Employed	Return	Unemployed
Age	26.10	26.37	25.78	25.57	34.47	26.23
	(10.72)	(11.34)	(9.94)	(10.35)	(12.11)	(16.21)
Education	9.64	9.57	9.72	9.73	8.19	9.38
	(2.80)	(2.68)	(2.93)	(2.74)	(3.36)	(2.66)
Health	1.71	1.73	1.69	1.71	1.80	2.08
	(0.70)	(0.72)	(0.69)	(0.70)	(0.71)	(0.86)
Wages	1397.10	1325.93	1485.32	1397.10		
	(957.28)	(797.62)	(1118.48)	(957.28)		
Marriage	0.28	0.30	0.26	0.26	0.64	0.31
	(0.45)	(0.46)	(0.44)	(0.44)	(0.48)	(0.48)
Minority	0.02	0.01	0.02	0.02	0.01	0.00
	(0.13)	(0.12)	(0.14)	(0.13)	(0.10)	(0.00)
Disabled	0.03	0.03	0.03	0.03	0.03	0.08
	(0.17)	(0.18)	(0.16)	(0.17)	(0.17)	(0.28)
Duration	25.99	18.94	34.14	10.45	230.63	336.96
	(68.64)	(52.18)	(83.05)	(21.39)	(86.11)	(232.05)
Hours	57.72	59.15	56.37	57.72		
	(15.99)	(16.21)	(14.81)	(15.99)		
Rent	179.48	188.17	169.42	182.22	141.24	140.75
	(145.71)	(168.39)	(113.22)	(150.10)	(26.07)	(91.55)
Transportation	32.45	30.70	34.49	32.51	32.85	22.85
	(27.53)	(31.11)	(22.55)	(28.36)	(7.08)	(23.64)
Networks	0.54			0.55	0.24	0.77
	(0.50)			(0.50)	(0.43)	(0.44)
Observation	1634	877	757	1525	96	13

1. Hours are working hours per week.
2. Health is a self-report variable which is scaled from 1 to 5. 1 is the best and 5 is the worst.
3. Cause describes reasons to exit the current state: d=1 if immigrants find a job and d=2 if immigrants return to their hometown
4. Wage, rent, transportation are monthly. The unit of duration is day.
5. Standard deviations are reported in parentheses.

estimates for those without networks.

Usually, male migrants and female migrants have different performances in the labor market. Tables 1.4 and 1.5 give the statistics summary based on gender. Table 1.4 shows information of male migrants. We can find the age, education level and marriage ratio are quite similar for unemployed male migrants as compared to employed male migrants. Female unemployed migrants are the oldest, have the lowest level of education and the highest marriage rate. The returned migrants, both male and female, have lower education level and fewer networks. Male migrants without networks have higher wages than those with networks, but for females, there are not significant differences in wages between the group with networks and the group without networks.

Table 1.5: Sample Summary for Female

	Female	Network		Labor Market Outcomes		
		1	0	Employed	Return	Unemployed
Age	26.08	26.57	25.43	25.53	30.70	37.09
	(9.94)	(10.46)	(9.18)	(9.55)	(10.86)	(13.24)
Education	9.21	8.93	9.58	9.35	7.97	6.76
	(3.23)	(3.15)	(3.31)	(3.21)	(2.84)	(3.48)
Health	1.81	1.83	1.77	1.80	1.76	2.09
	(0.72)	(0.74)	(0.70)	(0.72)	(0.64)	(0.87)
Wage	1153.57	1133.53	1181.63	1153.57		
	(521.40)	(517.55)	(525.93)	(521.40)		
Marriage	0.37	0.40	0.33	0.35	0.58	0.79
	(0.48)	(0.49)	(0.47)	(0.48)	(0.50)	(0.41)
Minority	0.02	0.02	0.01	0.02	0.00	0.03
	(0.13)	(0.14)	(0.12)	(0.13)	(0.00)	(0.17)
Disabled	0.02	0.02	0.03	0.02	0.01	0.03
	(0.14)	(0.13)	(0.16)	(0.15)	(0.11)	(0.17)
Duration	30.15	23.17	39.22	10.01	253.29	310.60
	(80.38)	(71.96)	(89.40)	(24.91)	(94.35)	(151.12)
Hours	57.50	59.25	55.54	57.50		
	(16.15)	(16.47)	(14.83)	(16.15)		
Rent	178.42	183.57	171.73	180.55	150.80	158.14
	(145.45)	(164.44)	(116.05)	(147.45)	(77.84)	(178.19)
Transportation	31.14	28.72	34.29	31.12	34.82	23.42
	(25.84)	(26.46)	(24.68)	(26.35)	(8.53)	(30.39)
Networks	0.56			0.58	0.23	0.62
	(0.50)			(0.49)	(0.42)	(0.49)
Observation	1349	826	636	1349	79	34

1. Hours are working hours per week.
2. Health is a self-reported variable which is scaled from 1 to 5. 1 is the best and 5 is the worst.
3. Cause describes reasons to exit the current state: d=1 if immigrants find a job and d=2 if immigrants return to their hometown
4. Wage, rent, transportation are monthly. The unit of duration is day.
5. Standard deviations are reported in parentheses.

1.5 Estimation Results

In this section, we give the estimation results of structural model and reduced form competing risk model. The competing risk model is the standard Mixed Proportion Hazard method. Since in the structural model, I assume the random variables: wage and the value of return migration are independent, I choose to estimate the reduced form competing risk model with independent unobserved variables. It will help us to compare the results.

1.5.1 Reduced Form

Table 1.6 shows the reduced form results which are based on the MPH competing risk model with Weibull distributions. The factor of age decreases the likelihood of accepting jobs but it does not have a significant effect on the likelihood of returning. Male are more likely to return migration. However, male migrants does not get employed faster than female migrants. This finding coincides with the structural results, which show female migrants have higher job arrival rates. Marriage decreases both of the likelihoods: married migrants are harder to find a job and also are not likely to return. As we expected, higher level of education helps migrants to increase the probability of finding jobs. If focusing on the returning, education does not have a significant effect on the decision of returning. Another interesting finding is relative to networks, the reduced form results show networks can significantly improve the probability of the transition from unemployment into employment. Also Networks can tie migrants to urban cities for a longer spell.

Table 1.6: MPH Competing Risk Model

	Employed		Returned	
	Coefficient	t value	Coefficient	t value
Age	-0.056	-3.941	0.045	0.891
Age^2	0.001	3.012	-0.001	-0.813
Male	0.006	0.351	0.505	3.454
Marriage	-0.154	-2.435	-0.089	-0.297
Minority	0.227	1.638	-0.109	-0.108
Education	0.020	2.856	-0.009	-0.240
Disabled	0.042	0.396	-2.089	-3.740
Networks	0.348	9.197	-0.753	-4.199
Constant	-0.700	-3.195	-15.949	-16.222
α	0.523	71.986	2.643	19.047
Num of Observation	3096			
loglike	-11537.689			

1.5.2 Structural Model

The structural estimation results in this paper are calculated based on the discrete distribution of values of return migration. We assume the discrete distribution with zero mean and evaluate the discrete distribution at three different points (η_1, η_2, η_3) with probability ($p_1, p_2,$ and, p_3). Heckman and Singer (1984) show this method will not hurt the consistency results even when we do not know the distribution shape of values of return migration.

Estimation procedure:

Estimation of the model is done by maximum likelihood using a combination of the Nelder-Meade simplex method and the BFGS algorithm to maximize the likelihood.

Table 1.7 presents the estimation results of structural model. Focusing on the effect of gender, we can find that female migrants receive more job offers than male migrants. When we calculate the partial effect in term of job arrival rates, male migrants have less than 0.0846% probability getting offers in one unit time (day). Although the effect is not big, it is still significant and if we consider in longer durations, it would have a larger effect on male migrants' job search. Both education and maturity in age could help migrants to receive more job offers but the effects are not significant. From the coefficient of disabled dummy variable, we can see that it is difficult for disabled migrants to receive job offers. They have close to 1% probability of getting a job offer. Looking at results for networks, we find that networks increase the probability of having more job offers by 0.55% points. The result shows networks seems to be a useful channel through which migrants can receive more offers.

Table 1.7 also presents the estimation results for the mean of wage distribution. Male migrants have higher wages on average. Age is a key variable in the wage distribution. Older migrants get offered higher wages on average. Higher levels of education also help migrants to achieve higher wages. Both results are consistent with standard human capital theory. Minority migrants do not have significant differences from majority migrants in term of wages on average. We also find the sign of minority migrants in job arrival rates are positive, although the number is not significant. The reason could be the migrants' jobs are homogeneous and there is no strong selection in terms of ethnicity. As with job arrival rates, we can see that disabled migrants also have lower wages on average. An interesting finding about networks is that migrants with networks have higher wages. This result is different from the unconditional results in Table 1.3. The structural model tells us the networks still have positive effects on wages. The results may be reflect the selection, i.e., the individuals who migrate and the individuals who have social networks. The next chapter deals with these selection problems by allowing for individuals invest in their social networks and by modelling individuals' migration decisions.

Table 1.7: Structural Estimation Result

	Job Arrival Rate	Wage Distribution	
Male	-0.274	Male	0.010
	(0.139)		(0.006)
Age	0.038	Age	0.015
	(0.022)		(0.003)
Health	0.002	Age^2	-0.000
	(0.002)		(0.000)
Education	0.862	Education	0.001
	(0.522)		(0.000)
Minority	0.276	Minority	-0.040
	(0.686)		(0.033)
Disabled	-2.841	Disabled	-0.078
	(0.849)		(0.032)
Networks	1.772	Networks	0.014
	(0.403)		(0.005)
Constant	0.055	Married	-0.043
	(0.015)		(0.016)
		Constant	-0.968
			(0.068)
		σ_w^2	1.107
			(0.032)
	Return Distribution	Search Cost	
α_1	-2.847	Transportation	0.023
	(1.036)		(0.011)
α_2	0.586	Rent	-0.031
	(0.050)		(0.009)
p_1	0.009	d_t	0.166
	(0.009)		(0.051)
p_2	0.126	d_r	-0.264
	(0.030)		(0.097)
		Constant	3.305
			(0.140)
Num of Observation	3096		
loglike	-162822.483		

1. Standard errors are given in parentheses.
2. Health is a self-reported variable which is scaled from 1 to 5. 1 is the best and 5 is the worst.
3. Cause describes reasons to exit the current state: d=1 if immigrants find a job and d=2 if immigrants return to their hometown
4. The unit of duration is day.

1.5.3 Comparison

In this section, I compare the results of the reduced form model with the structural model. MPH competing risk model can not distinguish the effects from the job arrival rates and from wage distribution. However, the job search model can separately identify these effects. For example, the MPH model shows that male do not have significant differences in terms of being employed in urban cities. However, the structural model shows that female migrants have higher job arrival rates but have lower wages on average.

Since the reduced form does not have the support from the specific structural model, I compare the unconditional density of each competing risk. From the data, I can construct the actual density of migrants' transition from unemployment to employment or return migration. I also can calculate the unconditional densities based on the parameters from the reduced form and structural model.

Table 1.8 presents a comparison between the reduced form model and structural model. The numbers in column of Day show the fraction of migrants who work or returned in a given duration. For example, the number of 0.728 in the column of data means 72.8% of migrants find jobs when they stay at urban cities for 10 days. Using the parameters of reduced form and structural models, we can calculate the unconditional densities based on different models. The cumulative distribution functions of the unconditional density are reported in Table 1.8. We can find the structural model fixes the data for employment state quite well. For example, the structural model shows 70.8% of migrants would find jobs when they spend at most 10 days in urban cities. Reduced form presents that 58.4% of migrants can find jobs in 10 days. The actual data tell us there are 72.8% of migrants who find jobs in 10 days. The reduced form model predicts the unconditional cumulative density function which is systemically lower than the actual cumulative density of migrants who exit to employment. But the performance of both models do not have a good job to fit the moment of returned migration behaviours. Generally speaking, the structural model has a better performance than the reduced form model.

1.6 Conclusion

The paper examines the job search process of rural migrants in China with the consideration of the option of return migration. Also, I analyse how social network affect migrants' labor market outcomes: job arrival rates and wages. The structural model in this paper follows the idea of Frijters and van der Klaauw (2006): I treat the transition from unemployment to return migration as a choice instead of a stochastic occurrence (i.e., Van den Berg (1990)). The rural migrants have the permanent option of returning when the value of continue search and expected value of employment is lower than the value of returning. Since the value of returning is a random variable, the reservation wages here are not constant values

Table 1.8: Ratio of Durations that end in Employment and Return Migration within a Given Duration

Work			
Day	Data	Reduced-Form	Structural
10	0.728	0.584	0.708
20	0.825	0.686	0.886
30	0.890	0.727	0.944
40	0.897	0.750	0.958
60	0.914	0.764	0.967
90	0.920	0.768	0.969
120	0.924	0.771	0.969
150	0.925	0.772	0.969
180	0.926	0.772	0.969
210	0.926	0.772	0.969
240	0.926	0.772	0.969
270	0.926	0.772	0.969
300	0.926	0.772	0.969
330	0.926	0.772	0.969
360	0.928	0.772	0.969
Return Migration			
Day	Data	Reduced-Form	Structural
30	0.000	0.000	0.000
60	0.000	0.000	0.000
90	0.006	0.000	0.000
120	0.009	0.001	0.000
150	0.015	0.001	0.000
180	0.020	0.001	0.000
210	0.022	0.001	0.000
240	0.029	0.001	0.000
270	0.031	0.002	0.000
300	0.045	0.002	0.000
330	0.052	0.002	0.000
360	0.056	0.002	0.000

1. The numbers in the column of data are actual ratios of people who end in the state of employment or return migration.
2. The numbers in the column 2 is the ratios which are calculated by the parameters in reduced form model.
3. Numbers in column 3 are calculated based on the parameters in structural model.

even though the model is still in the framework of stationarity.

Since the reservation wage cannot help to characterize the process of the model, I need to use the value of unemployment to characterize the transition rate and survival function. Hence, the theorem 1.3.1 shows that I can get the value of unemployment even without the information of reservation wage. The unique solution of Bellman equation makes it possible to estimate the whole structural model.

I use the new data source (RUMiC), which are national level data covering locations with more than 50% rural migrants in China. The data make it possible to capture the main story of rural migrants' job search process.

From the structural job search model, we can find female migrants can receive more job offers but have lower wages on average. Age and education have positive effects on both job arrival rates and the mean of wage distribution. Disabled migrants receive fewer job offers and also get lower wages on average. Minority migrants do not have significant differences from majority.

I also analyze how networks affect job arrival rates and wage distributions. As we expected, networks help rural migrants to receive more wage offers and also migrants with networks have higher wages on average.

Adding the option of return migration could help to model the selection of migrants who stay in cities. From the comparison of job search process between MPH competing risk model and the search model, we also can find search model has a better performance to fit the job search process in the data.

Bibliography

Banerjee, B. 1984. "Information Flow, Expectations and Job Search: Rural-to Urban Migration Process in India." *Journal of Development Economics* 15:239–257.

Banerjee, Biswajit. 1983. "Social Networks in the Migration Process: Empirical Evidence on Chain Migration in India." *Journal of Development Areas* 17:185–196.

Borjas, G. and B. Bratsberg. 1996. "Who leaves? The Outmigration of the Foreign-born." *Review of Economics Statistics* 78:165–176.

Borjas, G.J. 1999. *The Economic Analysis of Immigration.*, vol. 3A. Amsterdam: North-Holland.

Chiswick, B.R. 1978. "The effect of Americanization on the Earning of Foreign-born Men." *Journal of Political Economy* 86:897–921.

Flinn, C. and J. Heckman. 1982. *Models for the Analysis of Labour Force Dynamics*. Greenwich: JAI Press.

Frijters, P and Bas van der Klaauw. 2006. "Job Search with Nonparticiparion." *The Economic Journal* 116:45–83.

Galor, O. and O. Stark. 1990. "Migrants' Savings, the Probability of Return Migration, and Migrants." *International Economic Review* 31:463–467.

Hare, Denise and Shukai Zhao. 2000. "Lobor Migration as a Rural Development Strategy: a View from the Migration Origin." *Working paper* .

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

Heckman, James and B. Singer. 1984. "A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52:271–320.

Jasso, G. and M.R. Rosenzweig. 1982. "Estimating the Emigration Rates of Legal Immigrants Using Administrative and Survey Data: the 1971 Cohort of Immigrants to the US." *Demography* 19:279–290.

Kiefer, N.M. and G. Neuman. 1979. *Studies in the Economics of Search*, chap. Estimation of Wage Offer Distributions and Reservation Wages. Amsterdam: North-Holland.

Meng, Xin. 2000. *Labout Market Reform in China*. Cambridge University Press.

Mortensen, D.T. 1986. "Job Search and Labor Market Analysis." *Handbook of Labor Economics* 2:849–919.

Mortensen, D.T. and G.R. Neuman. 1984. *Choice or Chance? A Structural Interpretation of Individual Labor Market Dynamics*. Berlin: Springer-Verlag, 98–131.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labour Market." *Quarterly Journal of Economics* 118(2):549–599.

Murphy, Rachel. 1999. "Return Migration Entrepreneurs and Economic Diversification in Two Counties in South Jiangxu, China." *Journal of International Development* 11:662–672.

Van den Berg, G.J. 1990. "Search Behavior, Transitions to Nonparticiparion and the Duration of Unemployment." *Economic Journal* 100:842–65.

Zhang, Junfu and Zhong Zhao. 2011. "Social-Family Network and Self- Employment: Evidence from Temporary Rural-Urban Migrants in China." *working paper* .

Zhao, Yaohui. 2002. "Causes and Consequences of Return Migration: Recent Evidence from China." *Journal of Comparative Economics* 30:376–394.

———. 2003. "The Role of Migrant Networks in Labour Migration: the Case of China." *Contemporary Economic Policy* 21(4):500–511.

Chapter 2

Internal Migration with Social Networks in China

2.1 Introduction

A strong association between social networks and migration decisions has been consistently documented in numerous empirical studies. In most economic models, migration decisions are based on potential labor market outcomes. Social networks are often viewed as an important non-market institution through which individuals reduce market frictions and affect labor market outcomes. However, there are conflicting findings about the quantitative effects of social networks on labor market outcomes. For example, social networks may provide access to better jobs (Munshi, 2003; Edin, Fredriksson, and Aslund, 2003) or to less desirable ones (Borjas, 2000; Chiswick, Lee, and Miller (2005)). Although some researchers point out that individuals with social networks in destination places are more likely to migrate (eg., Munshi (2003)), there are not many papers which formally analyse how social networks affect individuals' migration behavior and their labor market outcomes.

In this paper, I construct a dynamic model with return and repeated migration, unemployment, and social network investment decisions that affect the evolution of social networks over time. The existing migration literature suggests two alternative mechanisms through which social networks may affect migration decisions and migrants' labor market outcomes. First, social networks may reduce migration costs (e.g., Carrington, Detragiache, and Vishwanath (1996); Munshi (2003)), decreasing individuals' migration reservation values causing individuals with networks to be more likely to migrate. Second, social networks provide information about labor markets and then increase the probability of getting job offers in the destinations (e.g., Montgomery (1991); Kono (2006); Goel and Lang (2012); Buchinsky, Gotlibovski, and Lifshitz (2012)). Under both of these mechanisms, individuals with social networks are more likely to migrate.

Although both of these mechanisms can explain why individuals with networks are more likely to migrate, they have different implications about migrants' earnings. Individuals with social networks have lower migration costs which cause to have lower reservation earnings. This means that migrants with networks are more likely to have lower earnings compared to similar individuals without networks. However, if social networks reduce search frictions, for example by increasing the job arrival rate, then migrants with networks will have higher earnings than similar individuals without networks. These different implications for migrants' earnings may be one reason for why some papers find positive earnings' effects while others find the opposite. The goal of this paper is to quantify the different roles that social networks may play with regard to labor market outcomes.

One issue concerning social networks is that they are unlikely to arise independently of individuals' labor market prospects. That is, individuals make investment choices in their social networks by comparing the loss from the payment of network investment to the benefit from increasing the probability of having a social network. In the literature, the common approach has been to look for natural or quasi-natural experiments as an attempt to deal with this problem.[1] In contrast, in this paper, I accounting for this possibility by formally modelling the social network investment decisions made by individuals. Modelling social networks with network investment decisions aids in our understanding of how individuals respond to market frictions through their social networks. Considering social network investment decisions also helps to evaluate potential government migration policies. The effects of government policies on market frictions and migration costs are likely to result in differential responses by individuals in terms of their social investment decisions and, ultimately, their migration outcomes. Failure to account for these feedback effects may lead to inaccurate policy evaluation.

Understanding different channels through which social networks operate is crucial for accurately designing migration policies. For example, the Chinese government aims to increase the urbanization rate to 60% by 2020, which means that an additional 100 million rural people will need migrate to urban areas.[2] Whether social networks are substitutes or complements to government policies aimed to increase migration may greatly affect their cost effectiveness.[3]

Besides accounting for the impact of social networks, the model I use in this paper also contains a number of mechanisms through which individuals' migration decisions are affected. First, I allow individuals to accumulate human capital

[1] Goel and Lang (2012) use a different-in-different approach to analyse how social networks affect labor outcomes, but they do not allow individuals to make their own decisions to invest in social networks.

[2] The urbanization rate is 53% by the end of 2014.

[3] The situation in China is obviously more complex than what is assumed in the model. Policies are implemented at both local and national government levels, and there exist some barriers to migration that are not formally included in the model. Despite the shortcomings, the model is able to showcase the importance of non-market institutions in developing countries (i.e., China).

within a search framework. Individuals' earnings reflect both their observed characteristics (e.g., education) as well as their location-specific human capital accumulation (i.e., urban and rural).

Second, individuals' earnings are also affected by frictions in the urban labor market. Individuals do not automatically have a job if they migrate. Instead, they need to search for one. Depending on the outcome of the search process, individuals may choose to stay in urban areas or return to rural areas. This setting incorporates one of the main features of rural-urban migration in China: most people do not migrate permanently.[4]

To study the role of social networks, this paper examines one of the largest migration episodes of the 20th century: rural to urban migration in China. The current internal rural-urban migration in China provides an ideal setting to examine the role of social networks in a labor market with frictions. Hare and Zhao (2000), Meng (2000) and Zhao (2003) show social networks are strongly correlated with rural-urban migration in China. Zhang and Zhao (2011) find social networks also affect migrants' subsequent labor market outcomes. However, these papers do not distinguish the social network effects through the two different channels discussed above.

The panel data I use for this study come from the Chinese Household Income Project (CHIP, 2007-2009). It is well suited to examine the effects of social networks on migration decisions and labor market transitions in China. First, the data cover most provinces of rural-urban migration in China. Second, the data contain enough information on social networks and labor market outcomes across different locations to identify the effect of social networks through migration costs and the job arrival rate. Finally, the data contain information on individuals' social network investment.

I estimate the model using the CHIP data. The estimation results show that social networks both significantly reduce migration costs and increase the job arrival rate. Individuals with networks have almost twice the job arrival rate compared to those without. Social networks reduce migration costs by 10% on average. To analyse the importance of these two channels, I simulate the model and show that migration decisions are affected more by the impact of social networks on reducing search frictions than by their impact on reducing migration costs. If I shut down the effects of social networks on both channels, only 14% of rural people migrate. Allowing social networks to only affect migration costs leads to 16% of rural people migrating. If social networks only increase the job arrival rate, 26% of rural people will migrate, compared to 29% in the data.

The simulation results also illustrate how individuals respond to the impact

[4]More than 45% rural migrants had the experience of return and repeat migration.

of social networks through network investment. When social networks affect both channels, 58% of individuals invest in their social networks. If social networks only lower migration costs, the fraction of individuals who invest decreases to 7%. When social networks only affect the job arrival rate, 53% of individuals invest in their social networks. The results also show that most individuals who invest in their social networks are the ones living in rural areas and the ones unemployed in urban areas.

I simulate three different policies to achieve the stated Chinese government's goal of a 60% urbanization rate: an unconditional lump sum subsidy for rural individuals who migrate, the provision of unemployment benefits for rural migrants in urban areas, and a migration cost subsidy for rural people, but only for those who have social networks in urban areas. The simulation results show that the policy of conditional lump-sum transfers for migrants will cost less than the other two policies. I also compare the effects of the policies to those obtained in a model estimated under the restriction that individuals do not invest in their social networks. I find that the government has to spend substantially more if individuals can invest in their social networks, as they will have an incentive to reduce their investment, partially counteracting the effects of the policies. That is, the government crowds out social network investment and as a consequence it will have to spend more, compared to the case of no investment responses.

The rest of the paper is organized as follows. Section 2 provides a review of the relevant literature. Section 3 presents background on rural-urban migration in China, describes the data in detail, and provides a preliminary empirical examination of the key mechanisms in the model. In Section 4, the model is described, identification conditions are provided, and I also describe the estimation procedure including challenges and solutions. Estimation results and counter-factual simulations are presented in Section 5. Section 6 concludes.

2.2 Literature Review

The existing migration literature has two main findings about the role of social networks. The first one is that individuals with social networks are more likely to migrate (i.e., Munshi (2003)). Hare and Zhao (2000), Meng (2000) and Zhao (2003) find social networks are also positively correlated with rural-urban migration in China. The second finding is that social networks affect migrants' labor market outcomes. For example, social networks may provide access to better jobs (Munshi (2003); Edin, Fredriksson, and Aslund (2003)) or to less desirable ones (Borjas (2000); Chiswick, Lee, and Miller (2005)). Zhang and Zhao (2011) examined the correlation between social-family networks and rural migrants' self-employment in China. They find social-family networks increase migrants' employment probability in urban areas.

Some of the literature on migration studies both out-migration and return migration, also known as circular migration. This literature is relevant to China, because most rural individuals engage in temporary migration. For example, my sample from the Chinese Household Income Project data (2007-2009) shows that more than 40% of rural individuals who have ever migrated to urban areas and have experienced return or repeated circular migration behaviours. Most empirical studies of rural-urban migration in China assume individuals have permanent migration. When analysing rural-urban migration in China, the prominence of circularity in behaviour of rural migrants makes this assumption undesirable.

Colussi (2006) also studies the role of social networks on migration behaviour in developing countries. He develops and estimates an equilibrium model to examine the impact of migrants' social networks on illegal Mexican immigration allowing for repeated circular migration. In his model, he assumes migrants' networks can increase the probability to find a job in U.S. In his model, however, social networks cannot affect migration costs. Individuals cannot invest in their social networks in his model, since the definition of social networks is the number of Mexicans from their home village.

Unlike Colussi (2006), Carrington, Detragiache, and Vishwanath (1996) consider the role of social networks on migration behaviour through the channel of migration costs. They build a dynamic model to analyse the phenomena that more black people migrated from the South to the North during the U.S. Great Migration period even though they faced a smaller wage gap. They claim that although the wage gap was larger before the 1930s, black people did not have social networks in the Northern part of the U.S. and migration costs were large. They show that social networks can influence individuals' migration decisions, since they may have lower migration costs if they have social networks in the destination place. However, they do not quantitatively examine how social networks affect migration costs, and assume that each individual has the same social network. They do not distinguish search frictions from migration costs in their model either.

Besides the friction of existing migration costs, search frictions in the destination labor markets will also affect individuals' migration decisions. Gemici (2011) compares migration behaviours between married couples and singles in a dynamic model of household migration with bargaining between family members. In her model, there exists uncertainty in the labor market and individuals' migration decisions are influenced by search frictions. She finds that migration of married couples occurs much more in response to the earnings of men than to the earnings of women, as women have lower wage offers, and a lower arrival rate of offers. Buchinsky, Gotlibovski, and Lifshitz (2012) examine the effect of a few alternative national migration policies on the regional location choices and labor market outcomes of migrant workers. In their paper, they estimate a dynamic discrete choice model that incorporates stochastic job offers and job terminations. However, these studies do not consider how social networks affect search frictions.

My study is also related to several papers which have analysed why social networks are correlated with labor market outcomes. Munshi (2003) follows Carrington, Detragiache, and Vishwanath (1996)'s idea to examine how social networks affect Mexican migrants to the U.S.. Since the size of social networks is endogenously determined, he uses last period rainfall as the instrument and finds that individuals are more likely to be employed and to hold a higher paying non-agricultural job when the size of network is exogenously larger. However, this study assumes that the probability of being employed in the destination is independent of the individuals' duration in the destination. This rules out the situation where individuals can invest in their networks and reduce search frictions.

In the theoretical literature, Kono (2006) shows that workers with social networks have fewer information deficiencies because they can use referral channels to find a job. Therefore, individuals with social networks will have higher wages than those without. Goel and Lang (2012) examine how social networks affect immigrants' labor market outcomes. In their model, the mechanism is that social networks can increase the probability to get a job offer. To avoid the endogenous problem of social networks, they employ the difference-in-difference approach.

Despite numerous empirical studies, it is not clear whether social networks have a positive effect on individuals' earnings. For example, social networks may provide access to better jobs (Munshi (2003); Edin, Fredriksson, and Aslund (2003)) or to less desirable ones (Borjas (2000); Chiswick, Lee, and Miller (2005)). The reason to an ambiguous effect on earnings is that social networks affect individuals' earnings through different channels: migration costs and search frictions. The net effect of social networks may vary in different economic environment.

2.3 Background, Data and Key Sample Statistics

2.3.1 Background of Rural-Urban Migration in China

Since 1958, the Chinese central government has restricted the mobility of the population. From 1958 to 1983, the rural people who had job offers in urban areas or recruitment letters from universities could migrate from rural to urban areas. Between 1984 and 1988, the central government did not restrict rural-urban migration. At that time, there is no market for exchanging food. People need to use food stamps to get food. However, if rural individuals migrated, they had to provide food stamps for themselves.[5] It was still hard for rural individuals to migrate since it was not easy to have enough food stamps to support themselves. This migration

[5]At that time, China was a planned economy. The amount of food for each individual was planned by the government. People needed to use food stamps to exchange food.

policy was suspended between 1989 and 1991. After 1992, the government began to encourage rural-urban migration and since 2000, the government started to reform the household registration system to encourage more rural individuals to migrate.[6] For example, in 2007, 12 provinces in China had cancelled the rural household registration, which means that rural individuals can have the same household registration as urban households in these provinces.[7] In these provinces, the local government does not distinguish between rural and urban residents any longer.

The easing of government restrictions on migration appears to have had a significant effect on people's migration decisions. Table 2.1 gives the inter-provincial migration in China from 1990 to 2005. There were 9.2 million people who migrated inter province between 1990 and 1995 and this number increased to 32 million between 1995 and 2000 and to 38 million between 2000 and 2005. Figure 2.1 gives the approximate number of rural migrants since 2000[8]. The number of rural migrants increases from 78 million to 145 million within 10 years.[9]

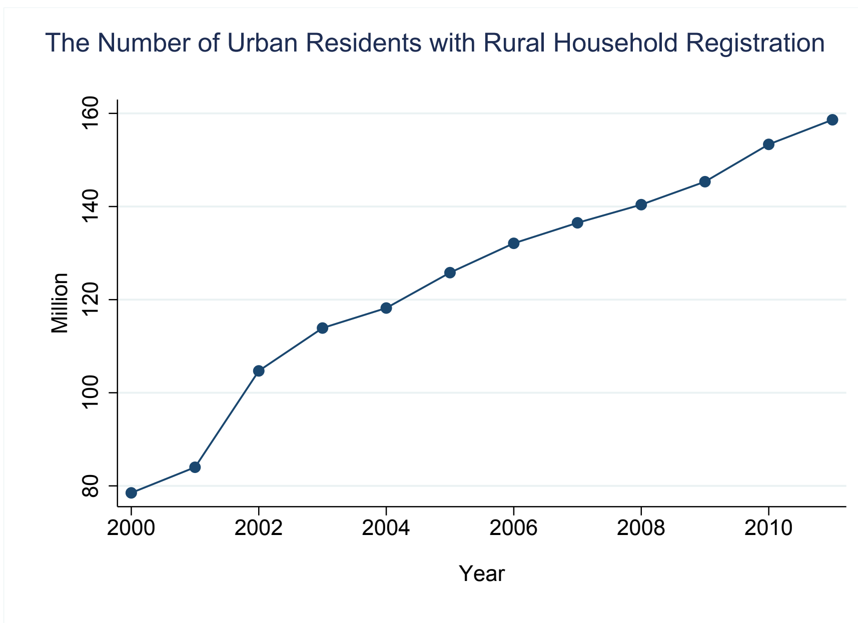


Figure 2.1: The Number of Stock Rural Migrants in China

[6]A household registration record officially identifies a person as a resident of an area and includes identifying information such as name, parents, spouse, and date of birth.

[7]These 12 provinces are Chongqing, Fujian, Guangxi, Hebei, Hubei, Hunan, Jiangsu, Liaoning, Shandong, Shanxi, Sichuan and Zhejiang.

[8]China Yearbook Rural Household Survey

[9]All numbers referred to the measure of the migrants' number is stock value in this paragraph.

Table 2.1: Interprovincial Migration in China, 1990-2005(In Thousands)

	1990-1995	migration				1995-2000		migration				2000-2005		migration			
		In	Out	Net	Net(%)			In	Out	Net	Net(%)			In	Out	Net	Net(%)
1	Guangdong	1947	221	1726	16.2	1	Guangdong	11501	438	11063	34.3	1	Guangdong	11996	1715	10281	27
2	Shanghai	726	122	604	5.7	2	Shanghai	2168	163	2005	6.2	2	Zhejiang	5062	1041	4021	10.6
3	Beijing	694	117	577	5.4	3	Zhejiang	2715	970	1745	5.4	3	Shanghai	3025	375	2650	7
4	Jiangsu	969	450	519	4.9	4	Beijing	1890	174	1715	5.3	4	Jiangsu	3290	1328	1963	5.2
5	Xinjiang	566	150	416	3.9	5	Xinjiang	1142	217	925	2.9	5	Beijing	2246	330	1916	5
6	Liaoning	435	197	239	2.2	6	Fujian	1346	625	722	2.2	6	Fujian	1934	802	1132	3
7	Tianjin	223	62	161	1.5	7	Jiangsu	1908	1241	667	2.1	7	Tianjin	908	107	802	2.1
8	Shandong	527	382	145	1.4	8	Tianjin	492	104	388	1.2	8	Xinjiang	577	182	395	1
9	Fujian	344	220	125	1.2	9	Liaoning	755	380	375	1.2	9	Liaoning	674	416	257	0.7
10	Hebei	503	417	87	0.8	10	Yunnan	733	398	335	1	10	Hainan	191	158	33	0.1
11	NeiMongol	275	249	27	0.3	11	Hainan	218	130	88	0.3	11	Ningxia	74	68	7	0
12	Shanxi	158	140	18	0.2	12	Shanxi	383	334	49	0.2	12	Tibet	26	31	-6	0
13	Tibet	38	28	10	0.1	13	Ningxia	129	87	41	0.1	13	Qinghai	74	85	-12	0
14	Hainan	104	102	2	0	14	Tibet	71	35	35	0.1	14	NeiMongol	394	417	-23	-0.1
15	Ningxia	49	54	-6	-0.1	15	Shandong	904	878	26	0.1	15	Yunnan	469	601	-132	-0.3
16	Qinghai	51	77	-25	-0.2	16	Qinghai	77	123	-46	-0.1	16	Shanxi	210	345	-135	-0.4
17	Yunnan	207	242	-35	-0.3	17	Hebei	770	872	-102	-0.3	17	Shandong	924	1123	-199	-0.5
18	Zhejiang	466	514	-49	-0.5	18	NeiMongol	325	441	-116	-0.4	18	Jilin	218	532	-315	-0.8
19	Shaanxi	163	265	-101	-1	19	Jilin	254	529	-275	-0.9	19	Gansu	118	494	-376	-1
20	Hubei	271	382	-111	-1	20	Shaanxi	423	719	-296	-0.9	20	Hebei	612	990	-378	-1
21	Gansu	140	251	-112	-1	21	Gansu	204	561	-357	-1.1	21	Shaanxi	255	827	-572	-1.5
22	Jilin	150	295	-145	-1.4	22	Heilongjiang	301	940	-639	-2	22	Heilongjiang	195	1020	-825	-2.2
23	Guizhou	152	402	-250	-2.3	23	Chongqing	448	1103	-655	-2	23	Chongqing	427	1437	-1010	-2.7
24	Jiangxi	125	514	-389	-3.6	24	Guizhou	261	1232	-970	-3	24	Guizhou	531	1766	-1235	-3.2
25	Heilongjiang	224	614	-389	-3.7	25	Guangxi	287	1838	-1551	-4.8	25	Guangxi	397	2123	-1726	-4.5
26	Guangxi	120	554	-434	-4.1	26	Hubei	606	2210	-1604	-5	26	Jiangxi	499	2476	-1977	-5.2
27	Henan	270	740	-470	-4.4	27	Henan	470	2309	-1839	-5.7	27	Hubei	501	2715	-2214	-5.8
28	Hunan	215	704	-489	-4.6	28	Jiangxi	236	2681	-2445	-7.6	28	Hunan	501	3328	-2827	-7.4
29	Anhui	155	744	-589	-5.5	29	Anhui	313	2893	-2579	-8	29	Henan	280	3433	-3154	-8.3
30	Sichuan*	395	1457	-1062	-10	30	Hunan	363	3261	-2899	-9	30	Anhui	671	3836	-3165	-8.3
						31	Sichuan	590	4396	-3806	-11.8	31	Sichuan	763	3941	-3178	-8.4
		9189	9189.00	0				32282	32282	0				38042	38042	0	

1. Before 2000, Chongqing was part of Sichuan province. The data for Sichuan from 1990 to 1995 includes Chongqing.
2. Source from NPSSO(1997), SC and NBS (2002,2007)
3. Net%=Net migration/Total national migration×100%.

After 2000, the central and local governments in China also proposed some policies to improve working and living conditions of rural migrants. For example, in early 2000, several provinces and cities such as Guangdong, Beijing, Shanghai and Xiamen started to set up social security schemes to cover rural labour migrants. A document issued by the State Council in May 2001 stated that local governments should provide nine years of compulsory education to migrant children through the public school system. Until the end of 2006, only a few local governments have actually implemented this policy (Liang (2006)). Although the central and local governments in China tried to change the rural household system and the associated discrimination, Chan (2012) states that the effects of those policies have not been large.

The government's migration policy may affect individuals' location choices. In the CHIP data, individuals in different cohorts show different migration patterns. Figure 2.2 shows that the fraction of individuals who migrate to urban areas from 2007 to 2009 is linearly increasing across different cohorts. Figure 2.3 examines the average ages at first migration across different cohorts. It shows a clear pattern that average age at first migration is decreasing linearly with cohorts. [10]

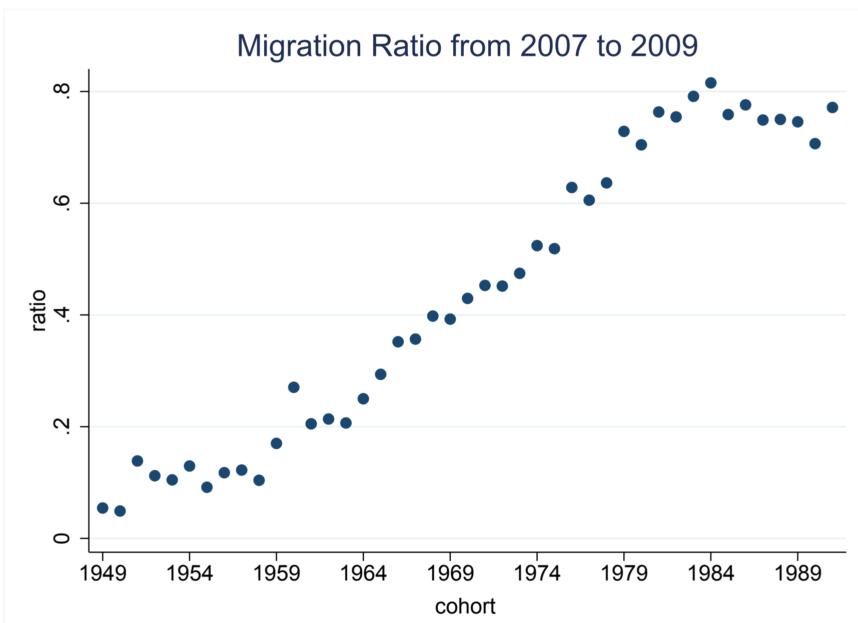


Figure 2.2: Migration Fraction by Cohort

[10]Both Figures 2.2 and 2.3 show that individuals in different cohorts have different migration patterns. In the structural model, I introduce the cohort effects in the migration cost function to incorporate the government policies' differential effects across cohorts experience.

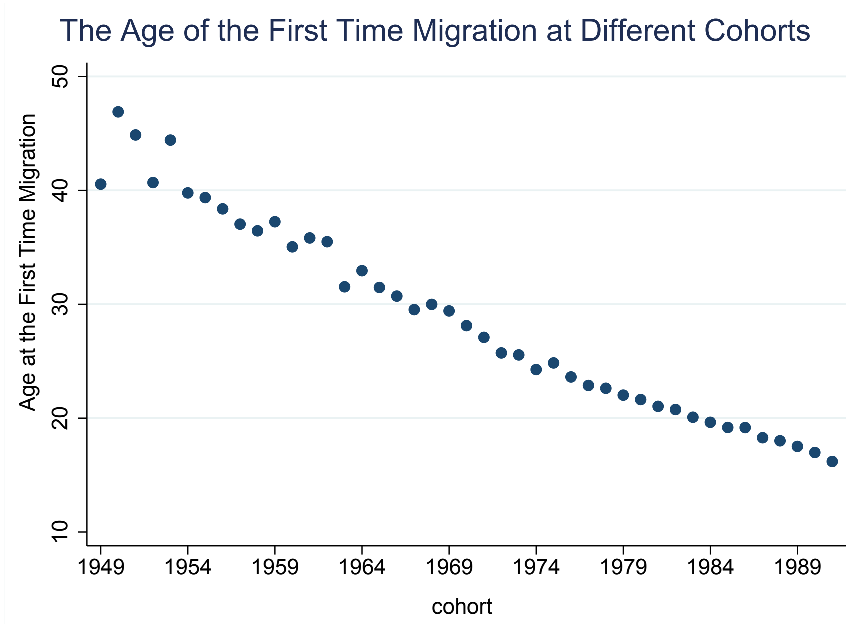


Figure 2.3: Average Age at the First Time of Migration in Different Cohorts

Next, since the government policies are changing over time, I examine whether a year effect is also an important factor. The survival analysis is used to see whether both cohort and the year effects are correlated with average age of individuals' first time migration. Table 2.2 gives the estimates assuming a loglogistic distribution. The coefficient for education shows that individuals with higher level of education will migrate earlier. The year dummies are the time when the central government made a large migration policy change. The year dummies (1984-1991) incorporate the policies for allowing migration but still with the need of food stamps, and the year dummies (1992-2000) incorporate the period of the transition between planned economy to market economy. These two year dummies are not significant at the 5% level. The year dummy (2001-2009) tries to capture the net effect of the policies made after 2000. The year dummy (2001-2009) is significant. Compared with cohort effects, year effects do not have a strong impact on individuals' migration choices.

2.3.2 Data

This study uses the first three waves (2007-2009) of the China Household Income Project (CHIP) panel survey.[11] This database is planned to be a five-year panel survey in China with the goal of studying issues such as the effect of rural-urban migration on income mobility and poverty alleviation, the state of education, and the health of children in migrating families.

[11]CHIP data (2007-2009) are part of the data of Rural Urban Migration in China.

Table 2.2: Survival Analysis of the Age of First Time Migration

	Coefficient	Standard Error
Education	-0.0016	0.0058
Cohort		
(1960-1969)	-0.5852	0.0364
(1970-1979)	-1.4855	0.0387
(1980-1991)	-2.4451	0.0350
Year		
(1984-1991)	0.1335	0.7340
(1992-2000)	0.1299	0.0842
(2001-2009)	-0.1654	0.0348
Constant	4.1304	0.0544
γ	0.5718	0.0077

1. The coefficient column is the estimation results of survival regression with loglogistics distribution.
2. The third column gives standard errors of estimates.
3. The omitted cohort dummy is the cohort (1949-1959).
4. The omitted cohort year dummy is the period before 1984 which is the period that rural-urban migration are prohibited.

Three representative samples of households were surveyed, including a sample of 8000 rural households, a sample of 8,000 rural migrant households, and a sample of 5,000 urban households in 9 provinces. The 9 provinces in the survey cover most provinces of the migration origin and destination in China. Figure 2.5 gives a map of the 9 provinces, which gives the net migration flow between 2000 and 2005. Table 2.1 shows that from 2000 to 2005, more than 68% of migrants moved into those 9 provinces while 52% of migrants moved out of those 9 provinces (NBS 2002, 2007).

Figure 2.4: Sample cities in China

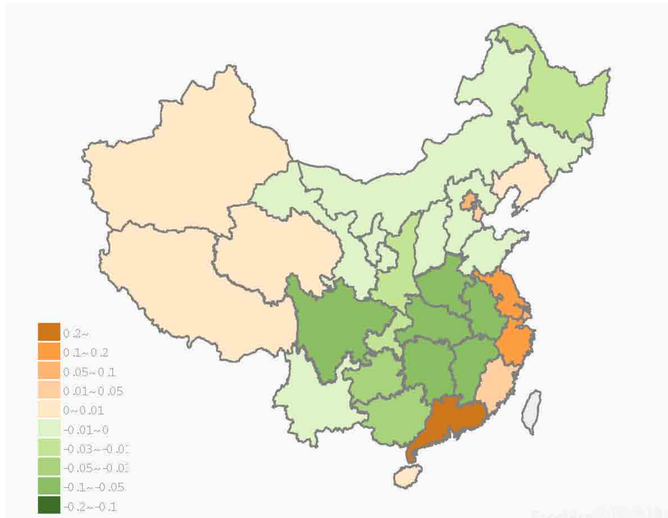


Figure 2.5: Migration flows in China from 2000 to 2005

In the analysis I use the CHIP rural household survey sample. Individuals in the rural sample are all born in rural areas and have rural household registration. Using the rural sample data, I can build the full history of the work experience for three years no matter where the individuals are located. The rural sample data include all individuals who have rural household registration. Individuals or their family members provide the information about the members in the household. For example, they provide the time when they leave their home, when they return and whether the destination is an urban county. Then, I can construct the monthly location history for each individual in the rural sample. The definition of migration is whether the urban residence location is out of his rural *hukou* (household registration) county.[12]

I analyse males in the rural sample for this study. I focus on males to avoid further expanding the model to take into account joint labor supply and fertility decisions. The samples contain information on work experience, job search durations, work locations, earnings, the presence of social network and social network investment decisions. Using this data, I can construct the location choices, job search durations, and work statuses for the individuals who are between 16 and 60 years old for the three year periods.

The total number of men for the 8000 households is 33,396 with 16,583 males. After restricting age between 16 in 2007 and 60 in 2009, the sample size shrinks to 11,385. In the data, there are 1030 observations missing the information on

[12]There are two main reasons why I do not use the migration sample. First, the response rate in the migration data is quite low. The attrition rate is above 70% for the three years of panel data. Second, I cannot follow the history of migrants' work experience using the migration samples. For example, migrants who return to their home towns are not surveyed in the migration sample.

fertility decisions or marital status. 1099 observations are missing their work experience information during 2007 to 2009. The sample used for estimation includes 9,256 males in 6400 households. The panel is balanced except for social network investment choices. Only the first two years' data contain the information about social network investment.

Social networks are defined as the presence of friends or relatives who are living in urban areas and are in contact with households.[13] Social network investment is whether they send monetary gifts to their friends or relatives. In the survey, people answer whether to give gifts to your friends or relatives and also report the monetary values of gifts. In the data, the gifts can be given to the friends or relatives who are living in rural areas. At the same time, individuals may build social networks through other channels (i.e., call each other, take care of friends' children or older family members). These two possibilities introduce measurement error in the social network investment. The estimation section provides details showing how I deal with this measurement error problem.

Table 2.3 displays selected descriptive statistics of the sample used in estimation. The variable of social networks is the presence of friends or relatives who living in urban area. More than 30% of individuals who live in households do not have social networks in urban areas. More than 60% of individuals invested in their social networks in 2007 and around 77% invested in 2008.

I restrict the sample to men who finished their formal education. The average years of education is 8.3 years. Since 1989, the central government has implemented the policy of 9 years' mandatory education in China, which is equivalent to completing middle school (or finish primary school). About 18% of individuals have less than 6 years education. Most people (i.e., 63%) complete middle school. Only 4% of individuals have post-secondary education.

I use the method proposed by Brandt and Holz (2006) to adjust earnings by location price index and the CPI price index. The base year is 2000. The average monthly earnings are around 1200 yuan and the average earnings in rural areas are less than 200 yuan.[14] The earnings in urban areas are six times the earnings in rural areas.

The definition of migration is whether the urban residence location is out of his rural *hukou* (household registration) county. Table 2.4 displays the descriptive statistics between migrants and non-migrants. First, migrants have higher education levels in general than non-migrants. The education levels are higher for individuals with networks than those without networks among both migrant

[13] In the CHIP data, social networks are measured at the household level. Individuals in the same household share the same social network.

[14] Yuan is the unit of Chinese currency. The exchange rate between Yuan and U.S. dollar was 8.28 in 2000.

Table 2.3: Summary Statistics for Estimation Sample

	Sample	observations
Fraction with networks	72.24%	9256
Network invest in 2007	61.26%	8026
	(0.49)	
Network invest in 2008	77.76%	7564
	(0.42)	
Education year	8.28	9256
	(2.14)	
Education level		9256
primary or less	18.34%	
middle school	63.49%	
high school	14.55%	
college or above	3.62%	
Urban earnings	1220.23	
	(1168.82)	
Rural earnings	153.09	
	(477.38)	

1. Earnings have been adjusted by location price index and cpi price index.
2. Earnings are in Chinese currency yuan which is closed to 1/8 of 1 U.S. Dollar in 2000.
3. Numbers in parentheses are standard deviations

and non-migrant groups. Second, migrants are much younger than non-migrants. The average age of migrants is 31. The average age of non-migrants is 42. The individuals with social networks are older than those without networks. Third, 60% of migrants get married whereas the fraction for non-migrants is 85%. Non-married individuals are more likely to migrate. Fourth, migrants with social networks have higher earnings than those without social networks. At the same time, migrants with networks have a smaller variance of the earnings. Non-migrants' earnings do not have significant differences between those with and without social networks. The average job search duration is 2 months. Migrants with networks have a slightly shorter job search duration than those without.

2.3.3 Preliminary Examination

Before introducing the structural model, I examine several correlations which are related to the key mechanisms proposed in this paper.[15] First, I document that there exists a strong correlation between social networks with both migration choices and subsequent labor outcomes.

[15]The analysis in this section uses the same data as the structural estimation in the next section.

Table 2.4: Descriptive Statistics for Migrants and Non-migrants

	Total	Migrants	Non-migrants
Education	8.28	8.63	8.01
	(2.14)	(1.92)	(2.25)
With Networks	8.37	8.73	8.14
	(2.12)	(1.93)	(2.22)
Without Networks	8.03	8.38	7.73
	(2.16)	(1.89)	(2.27)
Age	38.86	31.00	42.06
	(11.93)	(9.16)	(11.44)
With Networks	39.01	31.01	42.48
	(11.96)	(9.12)	(11.37)
Without Networks	38.40	30.97	40.90
	(11.84)	(9.28)	(11.55)
Marriage	0.78	0.61	0.85
	(0.41)	(0.49)	(0.35)
With Networks	0.79	0.61	0.86
	(0.41)	(0.49)	(0.35)
Without Networks	0.78	0.61	0.83
	(0.42)	(0.49)	(0.37)
Urban Earnings	1220.23	1220.23	
	(1168.82)	(1168.82)	
With Networks	1240.04	1240.04	
	(1232.21)	(1232.21)	
Without Networks	1162.12	1162.12	
	(909.06)	(909.06)	
Rural Earnings	153.09		153.09
	(477.38)		(477.38)
With Networks	158.81		158.81
	(480.50)		(480.50)
Without Networks	139.91		139.91
	(468.48)		(468.48)
Job Search Duration	2.20	2.20	
	(3.45)	(3.45)	
With Networks	2.19	2.19	
	(4.26)	(4.26)	
Without Networks	2.20	2.20	
	(2.99)	(2.99)	

1. Numbers in parentheses are standard deviations
2. Earnings have been adjusted by location price index and cpi price index
3. Job search period unit is monthly

Table 2.5 shows the correlation between social networks and migration choices. The second column gives the OLS regression results. After controlling for educa-

Table 2.5: The Relationship between Social Networks and Migration Decisions

Dependent Variable: Living in Urban Areas		
	OLS	Probit
Networks	0.0678	0.0420
	(0.0059)	(0.0060)
Education Year	-0.0045	-0.0026
	(0.0013)	(0.0013)
Married	0.0261	0.0012
	(0.0102)	(0.0095)
Number of Children	-0.0141	-0.0117
	(0.0034)	(0.0037)
Age	-0.0209	0.0019
	(0.0020)	(0.0021)
$Age^2 \times 100$	0.0037	-0.0242
	(0.0024)	(0.0026)
Constant	1.0771	0.3323
	(0.0370)	(0.0025)

1. The variable of social networks is the presence of social networks.
2. Married is the indicator variable for marital status.
3. The coefficients for the probit models (with and without instruments) are average marginal effects.
4. Numbers in parentheses are standard errors.

tion, marital status, the number of children and age, the coefficient of the presence of social network is significantly positive. The estimates in the Probit column display the average marginal effects on the probability of living in urban areas. It shows that the individuals with social networks are more likely to migrate. Regardless of functional form assumption, the correlation between social networks and migration choices is significantly positive. This finding is consistent with most migration literature: social networks are strongly correlated with migration behaviour.

Next, I examine the correlation between social networks and social network in-

Table 2.6: The Relationship between Social Networks and Network Investment

Dependent Variable: The Presence of Social Networks				
	OLS	2SLS	Probit	Probit with IV
Invest_{t-1}	0.0018	0.0151	0.0020	0.0358
	(0.0005)	(0.0040)	(0.0005)	(0.0134)
Education Year	0.0005	0.0005	0.0005	0.0007
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Networks_{t-1}	0.9732	0.9725	0.1126	0.1841
	(0.0005)	(0.0006)	(0.0019)	(0.0303)
Married	0.0033	0.0029	0.0037	0.0049
	(0.0007)	(0.0007)	(0.0007)	(0.0013)
Number of Children	-0.0012	-0.0011	-0.0013	-0.0019
	(0.0003)	(0.0003)	(0.0003)	(0.0006)
Constant	0.0144	0.0061	0.7606	0.7580
	(0.0012)	(0.0028)	(0.0002)	(0.0016)

1. The variable of social network investment is whether the individual send gifts to their friends or relatives at period $t-1$.
2. Married is the indicator variable for marital status.
3. In the third and fifth columns, the instruments for the 2SLS regression are the distance between rural county where the individual lives to Beijing, Shanghai and Guangzhou, which are top three cities in China.
4. The coefficients for the probit models (with and without instruments) are average marginal effects.
5. Numbers in parentheses are standard errors.

vestment. The second column of Table 2.6 presents the linear probability estimates. The coefficient on network investment is significantly positive even after controlling for education, last period networks, marital status and the number of children.[16]

Table 2.7 displays the results of employment regressions. The dependent variable is a dummy variable for being employed at the time of the survey. The results indicate that employment probabilities are positively correlated with the presence of social networks. Employment probabilities are also positively correlated with education. These findings are consistent with those in the literature on rural-urban migration in China (eg., Zhang and Zhao (2011)).

[16]The correlation remains even after I try to account for the potential endogeneity of social network investment, by using the distance between rural individuals' home town and Beijing, Shanghai and Guangzhou as instruments. The 2SLS column in Table 2.6 shows that this result is similar as the regressions estimated using a probit model. These results presented in the table 2.6 are not sensitive to the linear assumption.

Table 2.7: The Effects of Social Networks on Employment in Urban Areas

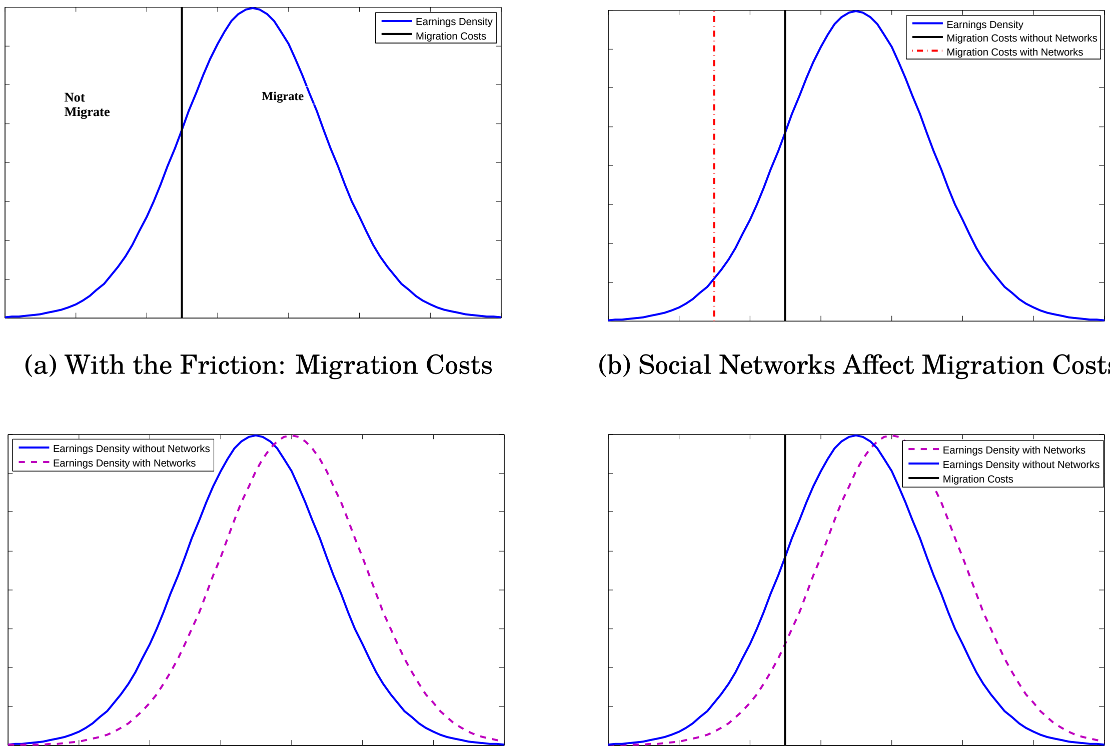
Dependent Variable: Employment		
	OLS	Probit
Networks	0.0279	0.0262
	(0.0021)	(0.0020)
Education Year	0.0013	0.0014
	(0.0005)	(0.0005)
Married	0.0116	0.0137
	(0.0028)	(0.0029)
Number of Children	-0.0095	-0.0099
	(0.0015)	(0.0092)
Age	0.0010	0.0009
	(0.0001)	(0.0001)
$Age^2 \times 100$	-0.0001	-0.0001
	(0.0000)	(0.0000)
Constant	0.6916	0.9151
	(0.0129)	(0.0009)

1. The variable of social networks is the presence of social networks.
2. Married is the indicator variable for marital status.
3. Numbers in parentheses are standard errors.

The regression results for migration behaviours and labor market outcomes suggest that social networks are an important determinant for migration choices and labor market outcomes. In the next section, I present a model that allows to quantify how important they are, identifying the mechanisms through which they operate, and illustrating how individuals' investment decisions shape the pattern of social networks and migration.

2.4 Model

I model individuals' migration decisions, labor market transitions and social network investment in a finite-horizon framework. To account for uncertainty regarding job offers, the migration decision problem is incorporated into a dynamic discrete time search framework. The decision period in this paper is taken to be one month in length. In each period, men receive flow utility associated with their current locations, incur moving costs if they decide to change their locations, and pay a cost if they invest in their social networks.



(a) With the Friction: Migration Costs (b) Social Networks Affect Migration Costs

(c) With the Friction: Job Arrival Rate (d) Social Networks Affect Job Arrival Rate

Figure 2.6: Illustration of How Social Networks Affect Migrants' Earnings

Individuals' flow utility comes from their earnings, unemployment benefits, psychic values of living in their home town, and a choice specific shock. The lifetime utility is given by current utility flow and the discounted stream of expected future utilities. Uncertainty comes from migration costs, search frictions in the urban labor market, the transition of earnings, the evolution of social networks and idiosyncratic shock to utility.

Before presenting the model, I illustrate why it is important to consider the impact of social networks through both migration costs and search frictions in a simple static example. Figure 2.6(a) gives the expected urban earnings' distribution for rural individuals. The solid black line is migration cost net rural earnings. The individuals whose urban earnings are larger than the migration cost would migrate. Figure 2.6(b) illustrates how the problem would look like when some individuals have social networks and these only affect migration costs. In this case, the migrants with networks would have lower migration costs and have lower reservation earnings. That is, since their migration costs are lower, they are willing to migrate for lower earnings than those without networks. Therefore, we would observe that migrants with social networks have lower earnings compared to migrants without social networks.

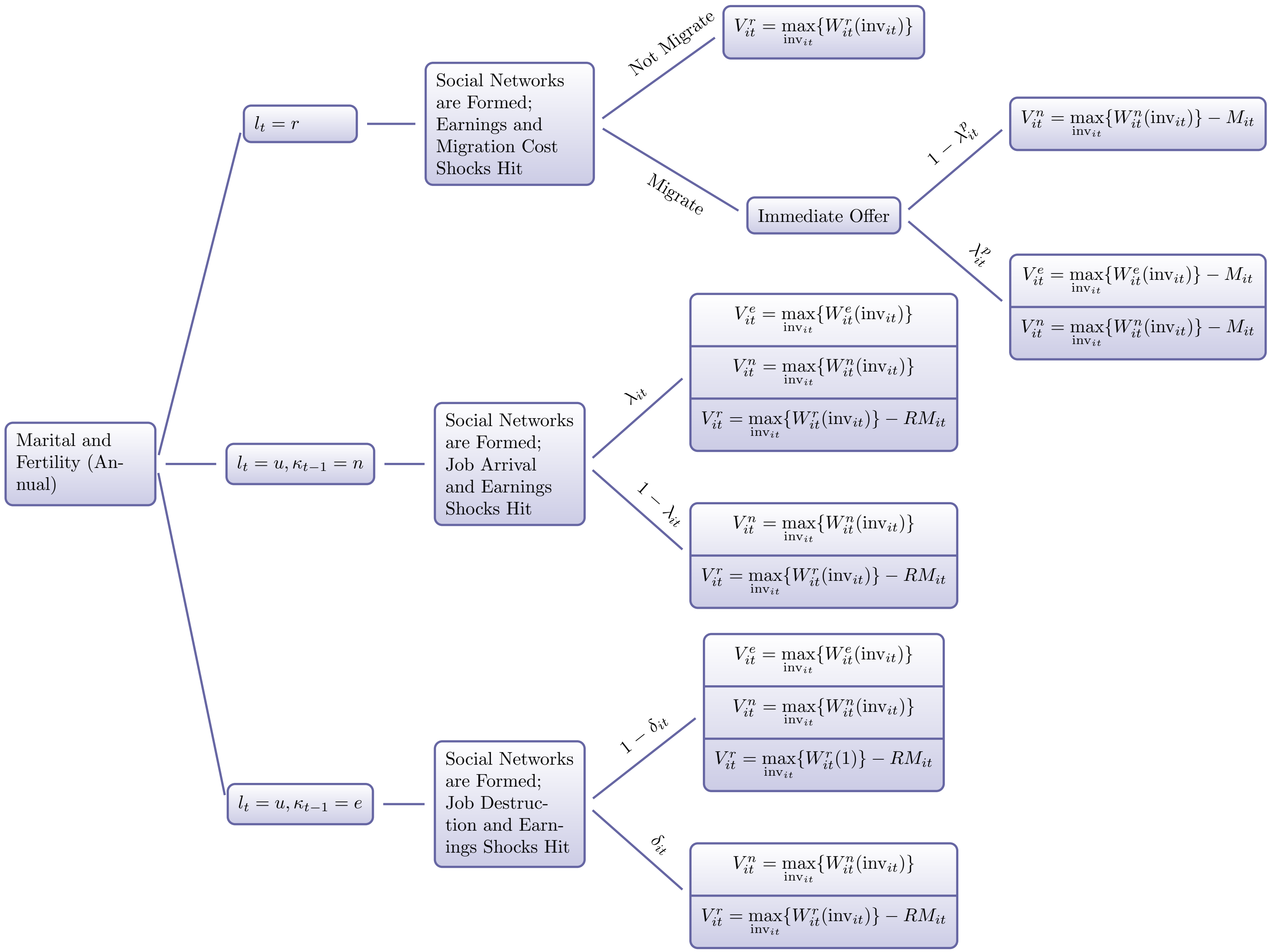
Figures 2.6(c,d) illustrate how this problem would look like when social networks only affect the job arrival rate. Figure 2.6 shows that individuals with social networks have a right-shifted expected earnings' distribution. Since in this case social networks only affect the job arrival rate, individuals have the same migration cost. Figure 2.6(d) shows that migrants with social networks have higher earnings compared with ones without social networks. The reason is that higher job arrival rates increase the reservation values of taking job offers.

These two mechanisms have opposite predictions about the correlation between social networks and migrants' earnings. These opposite predictions are also consistent with empirical findings: Munshi (2003) and Edin, Fredriksson, and Aslund (2003) find individuals with social networks have better labor outcomes; Borjas (2000) and Chiswick, Lee, and Miller (2005) support that migrants with networks have less desirable ones.

2.4.1 Timing of the decisions

Individuals' choices are made sequentially and also based on their current locations. Here, before describing the model, I specify the timing of individuals' decisions in the model is presented in Figure 2.7 and is as follows:

1. Individuals draw the marital and fertility shocks; marital and fertility transitions are exogenously formed annually.
2. At the beginning of period t, the shock for social networks is realised and individuals observe their social networks for period t.
3. If individuals are unemployed in urban areas, job arrival (or immediate job offer) shocks are realised; if individuals are employed in urban areas, separation shocks are realised.
4. (a) If individuals are in rural areas, both rural and urban earnings' shocks, migration cost shocks, and unemployment benefit shocks are drawn.
   (b) If individuals are in urban areas, both rural and urban earnings' shocks, and unemployment benefit shocks are drawn.
5. Following all of these shocks, location, employment and network investment choices are made jointly.



l_t is the location at the beginning of period t. κ_{t-1} is employment status at period $t-1$. $W^j(\text{inv}_{it})$ $j \in \{e, n, r\}$ is the choice specific value given the network investment decision.

Figure 2.7: The Timing of the Model

Timing based on location

In this paper, individuals make decisions based on their locations. Let $W_{it}^{j}(\text{inv}_{it})$ be the value of state $j \in \{e, r, n\}$ (e: employment in the urban area, r: rural, n: unemployment in the urban area). Let inv_{it} be an indicator that takes value 1 if an individual invest in his social network, and 0 otherwise. Denote $V_{it}^{j} = \max\{W_{it}^{j}(1), W_{it}^{j}(0)\}$. I will describe the decision process separately:

Urban

If the individual is in an urban area, which means he has already migrated, his choices are based on the following conditions:

1. If he just arrived in an urban area, he may receive an immediate offer.

    (a) If he receives a job offer (with probability λ_{it}^{p}), he will choose between two options, i.e. unemployment in the urban area, or accept the job offer ($\max\{V_{it}^{n}, V_{it}^{e}\}$). The social network investment choice is made at the same time.

    (b) If he does not get a job offer (with probability $1-\lambda_{it}^{p}$), he will be unemployed in the urban area (V_{it}^{n}). The social network investment choice is made at the same time.[17]

2. If he has already stayed for one or more periods in the urban area, he gets a job offer with probability λ_{it}, which is affected by the presence of social networks $sn_{i,t}$.

    (a) If he gets a job offer, he will choose between three options: unemployment in the urban area, accept the job offer, or return migration ($\max\{V_{it}^{n}, V_{it}^{e}, V_{it}^{r} - RM_{it}\}$). RM_{it} represents return migration cost. The social network investment choice is made at the same time.

    (b) If he does not get a job offer (with probability $1 - \lambda_{it}$), he will select between two options: unemployment in the urban area or return migration ($\max\{V_{it}^{n}, V_{it}^{r} - RM_{it}\}$). The social network investment choice is made at the same time.

3. If the individual works in an urban area, the exogenous job separation shock may hit him.

    (a) If he exogenously separates with the current job (with probability δ_i), he will choose to search for a job in an urban area or to return migrate ($\max\{V_{it}^{n}, V_{it}^{r} - RM_{it}\}$). The social network investment choice is made jointly.

[17] In the data, there are above 30% of rural migrants whose job search duration is less than two weeks. Since the model period is one month, I introduce the immediate offer to match the job search duration in the data.

(b) If he does not exogenously separate with the current job (with probability $1-\delta_i$), he will choose between three options: keeping this job, quitting this job to unemployment in the urban area, or quitting this job and returning home. The social network investment choice is made jointly.

Rural

If he is in a rural area, his earnings ω_{it}^r are drawn from the distribution $G(\omega^r)$. Hence, he knows the value of living in the urban area (V_{it}^r) and the value of migration ($\lambda_{it}^p \max\{V_{it}^e, V_{it}^n\} + (1-\lambda_{it}^p)V_{it}^n - M_{it}$). Then the migration decision is made based on equation 2.1:

$$\text{Mig}_{it} = \begin{cases} 1 & \text{if} \quad \lambda_{it}^p \max\{V_{it}^e, V_{it}^n\} + (1-\lambda_{it}^p)V_{it}^n - M_{it} > V_{it}^r \\ 0 & \text{else} \end{cases} \tag{2.1}$$

where λ_{it}^p is the probability of getting an immediate offer. Social network investment choice is made jointly.

2.4.2 Basic Structure

Earnings

Earnings are functions of education and work experience in rural and urban areas and location specific idiosyncratic shocks. The earnings of individual i in location $j \in \{u, r\}$ (u:urban, r:rural) at time t are described as

$$\omega_{it}^j = \beta_0^j + \beta_1^j \text{S}_i + \beta_2^j \text{exp}_{it}^r + \beta_3^j \text{exp}_{it}^u + \beta_4^j \text{exp}_{it}^{r\,2} + \beta_5^j \text{exp}_{it}^{u\,2} + \varepsilon_{it}^j \tag{2.2}$$

where S_i is years of education. In this paper, individuals accumulate their human capital through learning by doing via location-specific work experience (i.e., rural exp^r, urban exp^u). Here work experience in rural and urban areas are dependent on the history of endogenous decisions $\{d_{ik}\}_{k=1}^{t-1}$. The decisions include location, employment, and network investment choices. Shock terms ε_{it}^j are assumed i.i.d. across individuals, locations, and time, and they are normally distributed with mean zero and variance $\sigma_j^2, j \in \{u, r\}$ (u:urban, r:rural). Individuals know the current period transient components (i.e., ε_{it}^j), but they do not know values of future transient components. However, they do know the distribution of these future shocks and use them when taking expectations (i.e., rational expectations).

As specified in equation 2.2, social networks do not directly affect earnings. However, social networks may affect earnings indirectly through their effect on reservation earnings. On the one hand, individuals with social networks may have a higher job arrival rate which will increase their reservation values for accepting urban job offers. If so, individuals with social networks will have higher accepted earnings. On the other hand, social networks may reduce migration costs. Hence, for the same expected earnings, individuals with social networks are more likely to migrate because of lower migration costs. This implies that individuals with social networks may have lower reservation values of taking urban job offers compared with those without social networks. From the discussion above, the net relationship between social networks and migrants' earnings is not clear, since the effect of social networks goes through different channels (i.e., lower migration costs and higher job arrival rate).

Social Networks

In the model, individuals can invest in their social networks to strengthen connections with their friends (e.g., they may give gifts to their friends or contact with them by phone or mail). Social networks are formed according to the following dynamic probit model:

$$sn_{it} = \begin{cases} 1 & \text{if} \quad \beta_0^s + \beta_1^s \text{inv}_{it-1} + \beta_2^s \text{mar}_{it} + \beta_3^s \text{child}_{it} + \beta_4^s sn_{it-1} + \varepsilon_{it}^s > 0 \\ 0 & \text{else} \end{cases} \tag{2.3}$$

where sn_{it} is the indicator of social networks status at period t. It takes value 1 if the individual has social networks and 0 otherwise. inv_{it-1} is the individual's social network investment decision at period $t-1$. $\text{inv}_{it-1} = 1$ means that he invested in his social networks at period $t-1$, otherwise $\text{inv}_{it-1} = 0$. If the coefficient of β_1^S is positive, he can increase the probability of having social networks by the investment choice (e.g., giving gifts). mar_{it} is marital status at period t and child_{it} is the number of children at period t. The shock ε_{it}^s is i.i.d. across individuals and time. Individuals cannot observe future shock terms ε_{it}^s but they know the distribution of shocks.

In this model, the investment decision is a discrete choice that depends on the trade-off between the gain from increasing the probability of having social networks and the cost of investing. The key mechanism of investing in social networks is that it may increase the probability of having (or keeping) social networks, which may reduce migration costs and increase the job arrival rate. Individuals do not know their future shocks so their investment choices are based on their expectations of future shocks.

Migration and Return Migration Costs

If individuals migrate from rural to urban areas, they have to pay migration costs. One of the proposed channels through which social networks operate in the model is that they may affect migration costs directly.[18] Migration costs M_{it} depend

[18] Carrington, Detragiache, and Vishwanath (1996) build a dynamic macro model to examine the role of social networks on migration decisions. They also assume social networks reduce migration costs.

on the current period's presence of social networks, marital status, number of children, birth cohort and migration cost shock. I assume asymmetric migration costs: migration costs (M_{it}) may not be equal to return migration costs (RM_{it}). Migration and return migration costs in this paper are specified by the equations 2.4-2.5:

$$\mathrm{M}_{it} = \beta_0^m + \beta_1^m sn_{it} + \beta_2^m \mathrm{mar}_{it} + \beta_3^m \mathrm{child}_{it} + \beta_4^m \mathrm{cohort}_i + \varepsilon_{it}^m \tag{2.4}$$

$$\mathrm{RM}_{it} = \beta_0^{rm} + \beta_1^{rm} \mathrm{mar}_{it} + \beta_2^{rm} \mathrm{child}_{it} + \beta_3^{rm} \mathrm{cohort}_i \tag{2.5}$$

where sn_{it} is the indicator of the presence of social networks in urban areas at period t.

I also allow different cohorts to have different migration costs. This is to accommodate the fact that rural individuals across different cohorts have different migration patterns. As Table 2.2 shows, younger cohorts more likely to migrate earlier than older cohorts. the cohort term is an attempt to capture the net effect of the change of migration polices over four decades in China.[19]

Job Arrival and Destruction Rates

In the period in which people migrate to urban areas, they have to search for jobs from the unemployment state. Social networks may help individuals reduce search frictions in urban areas. The probability of getting an immediate offer λ_{it}^p upon arrival to the urban area, and the job arrival rate λ_{it} in urban areas are parametrised as:

$$\lambda_{it}^p = \frac{\exp\{\beta_0^{lp} + \beta_1^{lp} 1_{sn_{it}} + \beta_2^{lp} \mathrm{S}_i\}}{1 + \exp\{\beta_0^{lp} + \beta_1^{lp} 1_{sn_{it}} + \beta_2^{lp} \mathrm{S}_i\}} \tag{2.6}$$

$$\lambda_{it} = \frac{\exp\{\beta_0^{l} + \beta_1^{l} 1_{sn_{it}} + \beta_2^{l} \mathrm{S}_i\}}{1 + \exp\{\beta_0^{l} + \beta_1^{l} 1_{sn_{it}} + \beta_2^{l} \mathrm{S}_i\}} \tag{2.7}$$

To model exogenous job separation, the job destruction rate is parametrised as:

$$\delta_{it} = \frac{\exp\{\beta_0^{\delta} + \beta_1^{\delta} \mathrm{S}_i\}}{1 + \exp\{\beta_0^{\delta} + \beta_1^{\delta} \mathrm{S}_i\}} \tag{2.8}$$

[19] I model the cohort effect as linear. Figure 2.3 shows the average age of first migration across different cohorts. There is a linear relationship between cohort and the average age of first migration.

Marriage and Fertility Transition Process

I model annual marriage and fertility transitions with an exogenous process. The marital transition process is modelled as a continuous duration model with loglogistic distribution.[20] The survival function is:

$$\mathrm{Sur}_i(t) = (1 + (e^{-(\beta_0^{ma}+\beta_1^{ma} S_i)}t)^{1/\gamma})^{-1} \tag{2.9}$$

Here, $\mathrm{Sur}_i(t)$ is the probability of being single at age t, S_i is years of education, and γ is the parameter of the loglogistic distribution. Then, the conditional probability of getting married at period t is given by:

$$\Pr(\mathrm{mar}_{it} = 1|\mathrm{mar}_{it-1} = 0) = \frac{\mathrm{Sur}_i(t) - \mathrm{Sur}_i(t+1)}{\mathrm{Sur}_i(t)}$$

Fertility is determined by the following equation:

$$F_{it} = \begin{cases} 1 & \text{if} \quad \beta_0^f + \beta_1^f \mathrm{age}_{it} + \beta_2^f \mathrm{age}_{it}^2 + \beta_3^f \mathrm{child}_{it} + \beta_4^f \mathrm{child}_{it}^2 + \beta_5^f \mathrm{S}_i + \beta_6^f \mathrm{mar}_{it} + \varepsilon_{it}^f > 0 \\ 0 & \text{else} \end{cases} \tag{2.10}$$

Equation (2.16) shows that fertility is correlated with age, the number of children, marital status and education.

The Flow Value of Living in Rural Areas

The per-period utility in rural areas for individual i, at time t is given by

$$u_{it}^r = \omega_{it}^r - \nu \mathrm{inv}_{it} + \phi_{it} \tag{2.11}$$

where ω_{it}^r is rural earnings for individual i at time t, ν is the cost if he invests in his social networks at time t and ϕ_{it} is the psychic value of living in the rural area, which is given by

$$\phi_{it} = \beta_0^\phi + \beta_1^\phi \mathrm{age}_{it} + \beta_2^\phi \mathrm{age}_{it}^2 + \beta_3^\phi \mathrm{mar}_{it} + \beta_4^\phi \mathrm{child}_{it} \tag{2.12}$$

The reason I introduce a psychic value of living in rural areas is that, as documented in the migration literature (e.g., Kennan and Walker (2011)) people seem to place an additional value to their home towns, especially older individuals.

The Flow Value of Living in Urban Areas

Unemployment State

The per-period utility of being unemployed in urban areas for individual i, at time t is given by

[20] I assume there is no divorce for this marital transition. The annual divorce rate in rural areas is lower than 0.1%.

$$u_{it}^{n} = \xi_{it} - \nu \text{inv}_{it} \tag{2.13}$$

where ξ_{it} is the per-period utility of being unemployed in urban areas, and ν is the cost of investing in social network investment. As shows in equation 2.14, the per-period utility of being unemployed (ξ_{it}) is assumed to be a function of individuals' age, marital status and number of children. This setting allows elder people may have difficulties assimilating to a new environment so they may have different valuations of being unemployed in urban areas. Marital status and the number of children reflect the net value for an individual of living with his family.

$$\xi_{it} = \beta_0^{\xi} + \beta_1^{\xi}\text{age}_{it} + \beta_2^{\xi}\text{age}_{it}^2 + \beta_3^{\xi}\text{mar}_{it} + \beta_4^{\xi}\text{child}_{it} + \varepsilon_{it}^{\xi} \tag{2.14}$$

The shock term ε_{it}^{ξ} are assumed i.i.d. across individuals and time.

Employment State

The per-period utility of being employed in urban areas for individual i, at time t is given by

$$u_{it}^{e} = \omega_{it}^{u} - \nu \text{inv}_{it} \tag{2.15}$$

where ω_{it}^{u} is urban earnings for individual i at time t.

State Space

The vector of state variables for individual i at time t is denoted as H_{it}. State variables for a given time t include age, years of education, marital status, number of children, accumulated work experience in rural and urban areas, the presence of social network, and social network investment at period $t-1$. Control variables include individuals' decisions (i.e. migration, employment in urban areas, employment in rural areas, unemployment in urban areas, return migration, and social network investment decisions).

I assume that the transition of state variables is Markovian, and denote its transition probability by $\Pr(H_{it+1}|H_{it}, D_{it})$.

The transition of social networks is given by a dynamic probit model. Work experience in rural and urban areas is determined by the action history $D_{it} = \{\{d_{it}^{k}\}_{t=1}^{T}\}$, where $\{d_{it}^{k}\}_{t=1}^{T}$. $k \in \{1, \cdots, 5\}$ (1: migrate, 2: employed in urban, 3: employed in rural, 4: unemployed in urban, and 5: return migrate)[21].

[21]Social network investment choice is made jointly with these choices.

2.4.3 Value Function

The choice specific Bellman equations for each of the three states are

$$
\begin{aligned}
W_{it}^{n}(\text{inv}_{it}; H_{it}) &= u_{it}^{n} + \frac{\lambda_{it}}{1+\rho} E(\max\{V_{it+1}^{n}(H_{it+1}), V_{it+1}^{e}(H_{it+1}), V_{it+1}^{r}(H_{it+1}) \\
& -RM_{it+1}\}|H_{it}) + \frac{1-\lambda_{it}}{1+\rho} E(\max\{V_{it+1}^{n}(H_{it+1}), V_{it+1}^{r}(H_{it+1}) \\
& -RM_{it+1}\}|H_{it}) \\
W_{it}^{e}(\text{inv}_{it}; H_{it}) &= u_{it}^{e} + \frac{\delta_{it}}{1+\rho} E(\max\{V_{it+1}^{n}(H_{it+1}), V_{it+1}^{r}(H_{it+1}) - RM_{it+1}\}|H_{it}) \\
& + \frac{1-\delta_{it}}{1+\rho} E(\max\{V_{it+1}^{e}(H_{it+1}), V_{it+1}^{r}(H_{it+1}) - RM_{it+1}, V_{it+1}^{n}(H_{it+1})\}|H_{it}) \\
W_{it}^{r}(\text{inv}_{it}; H_{it}) &= u_{it}^{r} + \frac{1}{1+\rho} E(\max\{V_{it+1}^{r}(H_{it+1}), \lambda_{it+1}^{p} \max\{V_{it+1}^{n}(H_{it+1}), V_{it+1}^{e}(H_{it+1})\} \\
& +(1-\lambda_{it+1}^{p}) V_{it+1}^{n}(H_{it+1}) - M_{it+1}(H_{it+1})\}|H_{it})
\end{aligned}
$$

Based on the choice specific Bellman equation, the value function is stated by the following equation:

$$
V_{it} = \begin{cases} \lambda_{it} \max\{V_{it}^{n}, V_{it}^{e}, V_{it}^{r} - RM_{it}\} + (1-\lambda_{it}) \max\{V_{it}^{n}, V_{it}^{r} - RM_{it}\} & \text{if unemployed in urban} \\ (1-\delta_{i}) \max\{V_{it}^{n}, V_{it}^{e}, V_{it}^{r} - RM_{it}\} + \delta_{t} \max\{V_{it}^{n}, V_{it}^{r} - RM_{it}\} & \text{if employed in urban} \\ \max\{V_{it}^{r}, \lambda_{it}^{p} \max\{V_{it}^{n}, V_{it}^{e}\} + (1-\lambda_{it}^{p}) V_{it}^{n} - M_{it}\} & \text{if in rural} \end{cases} \tag{2.16}
$$

where $W_{it}^{j}(\text{inv}_{it}; H_{it})$ is the value of state $j \in \{e, r, n\}$ given the social network investment decision (e: employment in urban areas, r: rural, n: unemployment in urban areas). To simplify notations, I denote $V_{it}^{j} = \max\{W_{it}^{j}(1), W_{it}^{j}(0)\}$, $j \in \{e, r, n\}$. u_{it}^{j}, $j \in \{e, r, n\}$ is the flow utility at different states. M_{it} is migration cost, RM_{it} is return migration cost. λ_{it}^{p} is the probability of taking an immediate offer. λ_{it} is the job arrival rate and δ_{i} is the job destruction rate.

2.4.4 Identification

The model is a partial equilibrium model. I assume that the offered earnings' distributions are log normal. Based on the log normality assumption, the variance term of the earnings' distributions can be identified since I observe the accepted earnings. The distribution of unemployment value shocks in urban cities is assumed to be normal. The variance term of unemployment value shock σ^{ξ} can be identified from the probability of return migration given the variances of earnings' distributions (i.e., σ^{r}, σ^{u}).

The job arrival rate λ_{it} can be identified by the unemployment durations in urban areas. The probability of getting an immediate offer λ_{it}^{p} can be identified from the fraction accepting a job offer when they just arrive in urban areas. Since the fraction of employed migrants who switch into unemployment can be observed, the job separation rate δ_i can be identified from the behaviour of leaving high salary urban jobs to the unemployment state or back to home locations.

The psychic value ϕ_{it} of living in rural areas and migration costs M_{it} can be separately identified, because individuals only pay migration costs when they actually migrate while individuals receive the utility of psychic value every period. I set the monthly discount rate $\rho = 0.0025$ which gives 3% annual rate.

2.4.5 Objective Function and Solution Method

Individuals maximize their present discounted values of lifetime utility from the year in which they finish their education to a terminal age, $t = T$. Denote the utility associated with each choice as u_t^k. Then individuals make choices to maximise their objective function $V_{it}(H_{it})$:

$$V_{it}(H_{it}) = \max_{D_{it}^k} u_{it}^k(H_{it}) + \frac{1}{1+\rho} E(V_{it+1}(H_{it+1})|H_{it}) \tag{2.17}$$

The expectations operator E in equation (2.17) is taken with respect to the joint distribution of stochastic shocks $\varepsilon_{t+1}^{\xi}, \varepsilon_{t+1}^{u}, \varepsilon_{t+1}^{r}, \varepsilon_{t+1}^{m}, \varepsilon_{t+1}^{s}$,the probability of receiving a job offer, job separation, getting married, and having a child.

Given the finite horizon, the model is solved numerically through backward recursion of the Bellman equation. This procedure, however, cannot be applied directly due to the high dimensionality of the problem. Furthermore, the decision period in the model is a month which brings an additional computation burden. To reduce the computation burden associated with the high dimensionality of the problem, I adopt an approximation method similar to the one employed in Keane and Wolpin (1994). Instead of calculating continuation values at all points of the state space, I approximate them using a polynomial on the states. This is, at each t, I calculate the Emax functions (i.e., the continuation values) for a subset of the state space and estimate a regression function as a polynomial in those state space elements. I use the predicted values from the regression to approximate the alternative-specific value functions given by equation (2.17).

2.4.6 Estimation

Likelihood

The model is estimated by maximizing the likelihood function. For each individual, the data consist of the set of choices and outcomes:

- Choices: location, employment, and social network investment (i.e. $\{D_{it}^k : k = 1, \cdots, K\}$)
- Outcomes: earnings, presence of social networks

for all $t \in [t_{2007}, t_{2009}]$, where t_{2007} is individuals' age at the beginning of the year 2007 and t_{2009} is individuals' age at the end of the year 2009.

Let $c(t)$ denote the combination of choices (i.e., migration, employment, return migration and network investment) and outcomes at each period t. Let $t_{0i} \in [2007, 2009]$ denote the first period an individual is observed in the data. Notice that the state space includes both lagged variables (inv_{it-1}) and accumulated rural and urban work experience (expr_{it},expu_{it}). These variables, however, are nor observed for all individuals in the data. Let $\bar{H}_{0i}$ denote the value of the state space when an individual enters the sample. Then, if the probability of $\bar{H}_{0i} = \bar{h}_{t_{0i}}$ were known, the contribution to the likelihood for individual i would be:

$$\Pr(c(t_{0i}), \cdots, c(t_{Ti})|\bar{H}_{it}) = \sum_{\bar{H}_{0i} \in \Omega} \Pi_{t=t_{0i}}^{t_{Ti}} \Pr(c(t)|H_{it}, d_{it}) \Pr(\bar{H}_{t_{0i}}) \xi(p_m) \tag{2.18}$$

Equation 2.18 assumes that we know $\Pr(\bar{H}_{t_{0i}})$ and hence we can integrate it out. This, howerver, is not the case. It further assumes that the measurement of social network investment in the data corresponds exactly to the one in the model which is not the case either. In the next two sections, I describe a methodology to deal with these two problems.

Initial Condition Problem

As I discussed before, to calculate the likelihood for each individual, I need the state variables in the year they enter the sample. The data provide the whole marriage and fertility histories for each individual. However, it is missing the information on work experience in rural and urban areas for some individuals. I use a simulation method to solve this missing history problem. The basic idea is that, for the current value of the parameters, I simulate the transient shocks and use the value functions to simulate individuals' sequential decisions from an initial period until the time they enter the sample. I simulate such histories then to calculate the probability of observing an individual with $H_{t_{0i}} = h_{t_{0i}}$ to contribute the likelihood. The specific procedure is :

1. Given the current values of the parameters, the model is solved on grid and the value functions are saved.
2. Given the value functions, I draw from the distribution of the shocks to simulate a history from the time when finish formal education to t_{0i} for individual i. Let the value of the state at t_{0i} implied by simulation $r = 1, \cdots, R$ denoted by $\bar{H}_{t_{0i}}^r$.

3. Repeat step 2 R times[22]

4. Given $\{H^r_{t_{0i}}\}_{r=1}^R$, calculate $\Pr(H_{t_{0i}})$, which is what is needed to be calculated in the likelihood equation 2.18

Measurement Errors

As described in the data section, there is likely to exist measurement error for the social network investment variable. In the data, rural households only report whether they send gifts to their best 5 friends or relatives. These leads to two types of measurement errors: they may send gifts to someone outside of the best 5 friends or relatives, and(or) the gifts may be given to friends or relatives who are living in rural areas. In the estimation, I assume probabilities of having each type of measurement errors (p_m) are the same. Since I only observe the investment choices in 2007 and in 2008, the likelihood function for individual i is given by :

$$\begin{aligned}\Pr(c(t_{0i}), \cdots, c(t_{2009})|\bar{H}_{it}) &= \sum_{\bar{H}_{2007}\in\Omega} \Pr(\bar{H}_{2007})\Pi_{t=t_{2007.1}}^{t_{2007.12}} \Pr(c(t)|H_{it}, d_{it}) \\ &((1-p_m)^{m_{07}=a_{07}} p_m^{m_{07}\neq a_{07}})\Pi_{t=t_{2008.1}}^{t_{2008.12}} \Pr(c(t)|H_{it}, d_{it}) \qquad (2.19)\\ &((1-p_m)^{m_{08}=a_{08}} p_m^{m_{08}\neq a_{08}})\Pi_{t=t_{2009.1}}^{t_{2009.12}} \Pr(c(t)|H_{it}, d_{it})\end{aligned}$$

where $m_j, j \in \{2007, 2008\}$ is the model prediction of social network investment at period j, and $a_j, j \in \{2007, 2008\}$ is the data measure of social network investment at period j.

Estimation Procedure

The estimation algorithm is developed to incorporate both the initial condition problem, the measurement error problem, and to incorporate the assumed exogenous stochastic processes for marital status and fertility. The procedure is assumed as following:

1. Estimate the exogenous marital and fertility stochastic process and get the parameters Θ_1[23]

2. Given Θ_1, and the initial guess for the other parameters Θ_2, the model is solved on grids and the value functions are approximated as described in the section 2.4.5

3. For the individuals missing the value of the state, I draw the shock terms, and simulate their choices for the missing periods to calculate $\Pr(H_{t_{0i}})$ as described in how to solve initial condition problem

[22] I simulate 500 times for each individual who miss work experience information.
[23] Table C.1 gives the estimates of parameter Θ_1.

4. Calculate the likelihood and update the parameters Θ_2

5. Repeat from Step 2 to Step 5 until parameters Θ_2 converge

2.5 Empirical Application

2.5.1 Estimation

Estimates

The estimated parameters are reported in Tables 2.8. The parameter estimates are consistent with what one would expect. In particular, the effects of social networks on these two channels. First, the effect of social networks on migration costs is negative and significant. This means that the presence of social network reduces migration costs (i.e. they help migrants to settle down in urban areas). Second, social networks significantly increase the job arrival rate. The estimates are consistent with the idea that migrants can get job information from their friends or relatives in urban areas.

Table 2.9 displays the role of social networks through the two channels. The model estimates show that the average migration costs for individuals with networks are 89.2% of the value for those without networks. From the migration cost equation, we also find that married and individuals with children have larger migration costs. Also, the older cohorts have larger migration costs. Social networks may also affect search frictions. From the estimates of the job arrival rate, we see that social networks can reduce search frictions significantly. The average arrival rate for the individuals with networks ($\bar{\lambda} = 0.14$) is almost twice that for those without ($\bar{\lambda} = 0.07$).

The estimates of the earning equations in Table 2.8 show that the large gap between urban and rural areas does not mainly come from the returns to human capital. The difference mainly comes from the constant term. The return of education in urban areas is very low for rural migrants. In rural areas, the impact of an additional year of schooling on earnings (2%) is much higher than the one in urban areas (0.4%). The return of an additional month of urban work experience is higher than the return to an additional month of rural experience for urban areas. The opposite is also true. That is, rural experience return in rural areas is higher than urban experience return in rural areas.

From the unemployment equation, married and individuals with more children have a higher value on employment than single or childless individuals. At the same time, the estimates show that older people have a lower flow of utility values if they are in the unemployment state. The coefficients in the equation of psychic value of living in rural areas show that older people have higher flow utility of living

Table 2.8: Estimation Results

Earnings Equation(Urban)		Social Network Probit Equation	
edu year	0.0037	marriage	0.0690
	(0.0000)		(0.0141)
expu	0.0038	num of children	-0.0279
	(0.0000)		(0.0057)
expr	0.0036	sn_{t-1}	3.2374
	(0.0000)		(0.0162)
$expu^2 \times 100$	-0.0023	inv_{t-1}	0.6493
	(0.0000)		(0.2472)
$expr^2 \times 100$	0.0004	constant	-1.4204
	(0.0000)		(0.0127)
constant	6.8385	Psychic Value of Living in Rural Areas	
	(0.0014)	age	0.02638
Earnings Equation(Rural)			(0.0000)
		$age^2 \times 100$	0.0450
edu year	0.0208		(0.0000)
	(0.0002)	marriage	0.0021
expu	0.0015		(0.0199)
	(0.0000)	num of children	-0.0675
expr	0.0045		(0.0732)
	(0.0000)	constant	-0.0411
$expu^2 \times 100$	-0.0062		(0.0192)
	(0.0000)	Job arrival rate	
$expr^2 \times 100$	-0.0006	social network	0.7503
	(0.0000)		(0.0732)
constant	4.3097	edu year	0.0061
	(0.0013)		(0.0061)
Unemployment value		constant	-2.6004
			(0.0927)
marriage	0.1423	Job Destruction Rate	
	(0.1051)		
num of children	0.7132	edu year	0.0007
	(0.0392)		(0.0188)
age	0.0004	constant	-4.6729
	(0.0044)		(0.1344)
$age^2 \times 100$	-0.0071	Return Migration Cost	
	(0.0003)		
constant	0.0061	marriage	1.1140
	(0.1335)		(0.6391)
Migration cost		num of children	0.1409
			(0.1586)
social network	-1.3297	cohort	-0.01859
	(0.0630)		(0.1502)
marriage	0.1955	constant	8.0069
	(0.5719)		(6.6245)
num of children	0.0480	Immediate Offer Probability	
	(0.1753)	social network	-0.4516
cohort	-0.2795		(0.0524)
	(0.1510)	edu year	0.0162
constant	18.3926		(0.0092)
	(6.6705)	constant	1.5522
			(0.0926)

1. sn_{it} is an indicator variable that takes value 1 if the individual has social networks.
2. expu and expr stand for work experience in urban and rural areas respectively. They are both measured in months. Age is measured in years.
3. Cohort is defined by the birth year-1999.
4. Marriage is an indicator of marital status that takes value 1 if the individual is married.
6. Numbers in parentheses are standard errors.

Table 2.9: The Impact of Social Networks

	Key Estimates and Average Effects	
	Migration Cost	Job Arrival Rate
Social Networks Coefficient	-1.33	0.75
	(0.06)	(0.07)
Average		
With Networks	89.19%	0.14
Without Network	100.00%	0.07

1. I normalize the average migration costs for rural migrants without social networks in the model to 1. Hence, migration costs are presented as relative to 1, (i.e., 89.19% means that the average migration costs for individuals with networks is 89.19% for those without social networks.)
2. The first panel in the job arrival rate column presents the point estimate for social networks in the job arrival rate equation. The second panel shows the calculated average job arrival rate for rural migrants who are not employed depending on their network status.
3. Numbers in parentheses are standard errors.

in rural areas. This finding is consistent with the migration literature (e.g., Kennan and Walker (2011)) and the my data observation that migrants are younger than the individuals living in rural areas.

Model Fit

In this section, I evaluate the model fit, by using the estimates to simulate individuals' behaviours and compare the simulated results to the data moments.[24] Tables 2.10 and 2.11 give the comparison between the model predictions and data moments.[25] Table 2.10 shows the comparison to the earnings' moments. The data column gives the selected data moments for both migrants' and non-migrants' earnings including the mean and variance of log earnings. The other column gives the simulated moments based on the model estimates. Although the simulated standard deviations of log earnings are slightly larger than those in data, the calculated moments fit the data quite well. The model can successfully capture that earnings with networks are higher than the earnings without networks.

Table 2.11 gives the model fit for choices. The simulated moments can fit the fraction of the individuals with networks quite well. For example, in the data,

[24]For each individual, I simulate 100 times and the results reported are the mean of simulation results.

[25]I simulate the decisions for each individual from the age of finishing formal education to the age of 60. The moments calculated are based on the simulation results from the year 2007 to 2009.

Table 2.10: Model Fit: Earnings

	Data	Model
Migrants:		
Log(Earnings)	7.1068	7.1517
sd(log(Earnings))	0.2681	0.3273
with Networks	7.1229	7.1549
sd(log(Earnings))	0.2695	0.3232
without Networks	7.0580	7.1435
sd(log(Earnings))	0.2612	0.3376
Non-migrants:		
Log(Earnings)	5.0310	5.0257
sd(log(Earnings))	0.9683	1.1847
with Networks	5.0677	5.0283
sd(Earnings)	0.9547	1.1867
without Networks	4.9410	5.0190
sd(Earnings)	0.9989	1.1796

1. Migrants include people who currently work in urban areas.
2. sd(log(earnings)) stands for the standard deviation of log earnings.

Table 2.11: Model Fits: Choices

	Data	Model
Fraction of Individuals with Networks	0.7224	0.7317
Fraction of Rural Individuals Living in Urban Areas	0.2892	0.2734
Fraction with Networks	0.2179	0.2078
Fraction without Network	0.0712	0.0656
Fraction of Return Migrants	0.0075	0.0179
Fraction with Networks	0.0054	0.0142
Fraction without Networks	0.0022	0.0037
Fraction of Rural Individuals Migrating in a Given Month	0.0076	0.0142
Fraction with Networks	0.0054	0.0114
Fraction without Networks	0.0022	0.0027
Fraction of Individuals Getting Immediate Offer upon migrating	0.0038	0.0028
Fraction with Networks	0.0027	0.0025
Fraction without Networks	0.0011	0.0003
Average Job Search Duration (months)	2.1972	2.1059
Average Job Search Duration with Network at Initial Period	2.1929	2.0968
Average Job Search Duration without Network at Initial Period	2.1989	2.1458

1. The data column provides the moments calculated based on the observations during 2007-2009.
2. The numbers in the table are averages dividing the 36 months (2007-2009) observed in the data.

72.2% of individuals have social networks. The simulated moments are 73.2% in the model.

When examining the composition of the migrants, we find that both the model can capture rural individuals' migration choices quiet well. For example, in the data, 28.9% of rural individuals living in urban areas. When examining the decomposition of rural individuals living in urban areas, we can find the model also matches the moments really well. Among the 28.9% of rural individuals who live in urban areas, that 21.8% have social networks and 7.1% are without social networks. The model predicts this decomposition as 20.8% and 6.5%.

From Table 2.11, we can also see that the job search duration can be matched quite well: in the data, the average job search duration is 2.20 months; the model predicts the job search duration is 2.11 months. The model also captures the behaviours of accepting immediate offers after migrating well. For example, there are 0.3% of rural individuals who get a job immediately after migrating in the data. The model predicts the same number. As the table shows, the model also fits the moments related to return and repeat migration reasonably well.

2.5.2 Decomposition Analysis

In this section, I conduct counterfactual simulations to decompose the effects of social networks through migration costs and the job arrival rate. I then examine how social networks affect rural individuals' migration and social network investment choices.

To assess the effects of social networks on migration costs and labor market search frictions, I simulate the model under three different restrictions on the parameters. In the first specification, I turn off the effects of social networks on both channels, (i.e., $\beta_1^m = 0$, $\beta^{l^p_1} = 0$, $\beta_1^l = 0$). In this case, social networks play no role in the model. In the second specification, I turn off the effects of social networks on the job arrival rate (i.e., $\beta^{l^p_1} = 0$, $\beta_1^l = 0$). That is, social networks are only allowed to affect migration costs. In the third specification, I turn off the effects of social networks on migration costs (i.e., $\beta_1^m = 0$).

Table 2.12 presents the decomposition simulation results for the model in terms of the role of networks (i.e., neither of two channels, only affect migration costs, or networks only affect the job arrival rate). The second column presents the unrestricted (i.e., allowing social networks to affect both channels) predictions to use as a baseline case. The third column shows the model prediction if social networks do not affect either the job arrival rate or migration costs. Without the effect of social networks, only 14.0% of rural individuals will live in urban areas. The difference is more than 13% points when compare this to the results when network effects exist (27.3%). When social networks only affect migration costs (Column 4), 15.5% of rural individuals will live in urban areas. The job arrival rate column shows that

Table 2.12: Counterfactual Results: Social Networks

	Model	Social Networks Only Affect		
		Neither	Migration Cost	Job Arrival Rate
Fraction of Individuals with Networks	73.17%	71.05%	71.31%	72.43%
Fraction of Migrants	27.34%	13.97%	15.52%	26.63%
with Networks	20.78%	9.73%	11.35%	19.68%
without Networks	6.56%	4.24%	4.17%	6.94%
Fraction of Non-migrants	72.66%	86.03%	84.48%	73.37%
with Networks	52.40%	61.32%	59.96%	52.75%
without Networks	22.26%	24.71%	24.52%	20.62%
Fraction of Unemployed Migrants	4.93%	3.29%	3.90%	4.31%
Fraction of Network Investments	58.41%	0.00%	7.22%	53.10%

1. Migrants include people who were born in rural areas and resided in urban cities who can be employed or unemployed.
2. The column of model shows the benchmark values simulated by the model estimates.
3. The neither column gives the simulation results when social networks affect neither migration costs nor the job arrival rate.
4. The column of migration cost gives the counterfactual results if social networks only reduce migration costs.
5. The column of job arrival rate gives the counterfactual results if social networks only increase the job arrival rate.
6. The row of equalization monthly tax shows how much individuals would like to pay per month for their lives on average to achieve the same utility when not allowing them to invest in their social networks.

if social networks only increase the job arrival rate and have no effect on migration costs, 26.6% of rural individuals migrate. These simulation results show that the effect that social networks reduce search frictions is much more larger compared to reducing migration costs.

Table 2.12 also how individuals' social network investment choices also respond to the effect of social networks. For example, without the effects of networks, no one will invest in networks. When social networks only affect migration costs, about 7% of individuals will invest. When networks affect the job arrival rate, since the impact is even larger than the migration cost channel, more individuals choose to invest in their social networks.

Table 2.13 gives the decomposition results of the social network investment behaviours. The benchmark column provides the simulation results based on the estimates from the model. Individuals' investment decisions are affected by their locations and employment states. For example, when social networks affects both migration costs and the job arrival rate (i.e., in the benchmark column), 66.5% of rural individuals invest in their social networks and only 33.9% of employed rural migrants invest in social networks. The reason is that social networks only affect the continuation values when individuals are employed in urban areas. Therefore, employed rural migrants have less incentive to invest in their social networks. Under two restricted specifications, we still can find the pattern that most investors are still the individuals living in rural areas and the unemployed individuals in urban areas. The model results present that individuals effectively use and invest in social networks to optimize their migration decisions and labor market outcomes.

Table 2.13: The Responses of Social Network Investment

	Benchmark	Social Networks Only Affect Migration Costs	Social Networks Only Affect Job Arrival Rate
In Rural Areas			
Individuals who invest	66.53%	7.95%	68.54%
Individuals who do not invest	33.47%	92.05%	31.46%
In Urban Areas			
Employed			
Individuals who invest	33.93%	2.29%	30.87%
Individuals who do not invest	66.07%	97.71%	69.13%
Unemployed			
Individuals who invest	50.06%	6.15%	44.54%
Individuals who do not invest	49.94%	93.85%	55.46%

1. The benchmark uses the estimates from the model which allowing individuals to invest their social networks.
2. The column of migration cost gives the counterfactual results if social networks only reduce migration costs.
3. The column of job arrival rate gives the counterfactual results if social networks only increase the job arrival rate.

2.5.3 Policy Simulations

Since the Chinese government is trying to increase the urbanization rate to 60% by 2020, I propose three different policies all of which will achieve this aim. The policies include providing monthly unemployment benefits for rural migrants in urban areas, two types of lump-sum subsidies for migration costs.

First, I need to calculate the increase in the fraction of rural individuals migrating to urban areas to achieve the government's goal. Based on the annual report of National Bureau of Statistics of China (NBSC), an additional 118 million rural individuals will need to migrate to urban areas.[26] Since I only consider rural male individuals in my data, this translates to an additional 76 million rural males migrating.[27] Therefore, the fraction of total migrants will be about 50%.[28] When I

[26]This number is calculated based on total population in 2011. (i.e. $1347.7 \times 0.6 - 690.8 = 117.8$)

[27]This number is calculated by the total additional migrants times the fraction of male migrants (i.e. $118 \times 64.5\% = 76$). 64.5% is the fraction of male migrants. Since NBSC does not provide the number of rural migrants, I use the the number of rural men who did not migrate in 2011 divided by the fraction of rural individuals who reside in rural areas to calculate the number of rural men with rural registration (i.e. $\frac{250.59}{0.71} = 352.95$). I then calculate the number of male migrants (age 15-64) by subtracting the number of rural men who do not migrate, (i.e. $352.95 - 250.59 = 102.36$ million). China Yearbook Rural Household Survey states that the total rural migrants in 2011 is 158.6 million. The fraction of migrants who are male is $102.36/158.6 = 64.5\%$.

[28] The fraction is the sum of current rural male migrants and the additional rural males

simulate different policies, I target the fraction 50% of rural men migrate to urban areas in my data when I simulate the model.

Table 2.14: Policy Simulation Results: Government Budget

	With Network Investment		Without Network Investment	
	Policy (Yuan)	Government Budget (Billion Yuan)	Policy (Yuan)	Government Budget (Billion Yuan)
Unemployment Benefit	424	11.50	403	11.39
Unconditional Migration Cost	639	3.56	569	2.02
Conditional Migration Cost	279	1.10	101	0.49

1. In the model with network investment, individuals allow to invest their networks by send gifts to their friends or relatives.
2. The the model without network investment, social networks are formed by a dynamic probit process.
3. This table provides the three different policies' simulation results, all of which achieve the goal of urbanization rate (i.e., 60% by 2020)
4. The row of unemployment benefits shows the monthly value of unemployment benefits the government provides and the total budget the government pays to achieve the goal.
5. The row of unconditional migration cost gives the value of a lump sum subsidy for migration costs if rural individuals migrate and the total budget the government pays.
6. The row of conditional migration cost gives the value of a lump sum subsidy for migration costs if rural individuals with social networks in urban areas migrate and the total budget the government pays.

Table 2.14 provides the simulation results for the policy counterfactual simulations in three specifications. The first one is that the government provides monthly unemployment benefits in urban areas for rural migrants. This policy will increase the value of living in urban areas and decrease the return migration. The second policy is an unconditional lump sum subsidy for migration costs when rural individuals migrate to urban areas. The third policy provides a conditional lump sum subsidy for migration costs if migrants have social networks in urban areas.[29]

In Table 2.14, the second and third columns give the specifications of three policies when allowing for individuals' social network investment. If the government implements any of these policies, the urbanization goal can be achieved. The value of 424 means that the government pay 424 yuan of unemployment benefits in urban areas monthly; The value of 639 in the row of migration cost means the policy of a lump sum subsidy for migration costs (639 yuan) per person, if rural individuals migrate to urban areas. The value of 279 in the conditional migration costs row denotes the policy of providing lump-sum subsidy for migrants who have social networks (279 yuan) per person. Under these three policies, the goal of urbanization rate can be achieved. I then can compare these policies in term of government

who will migrate divided by the total number of individuals with rural household registration, (i.e.(76.0+102.4)/353.0=50.5%).

[29]There exists similar migration policy in Canada. If the individual has relatives or family member in Canada, he will be more easily to pass the immigration requirements from the Canadian government.

budgets. The government will spend less to implement the both policies of migration cost subsidy than the unemployment benefit policy to reach the goal of urbanization rate (2.02 (unconditional subsidy), 1.10 (conditional subsidy) v.s. 11.39 (unemployment benefit) billion Yuan). Since the conditional subsidy policy encourages individuals to invest their social networks, the government will spend less compared to the unconditional migration cost subsidy.[30]

Table 2.14 also provides the policy simulations for the model without social network investment decisions. Comparing the models with and without social network investment decisions, we find the government needs to spend more to attract rural people to migrate in the model with network investment decisions. The reason is that individuals will invest less in social networks or not at all when the government tries to reduce migration costs or search frictions in urban areas. To offset the impact of less network investment, the government has to spend more (i.e. 11.50 v.s. 11.39 for unemployment benefits; 3.56 v.s. 2.02 for unconditional migration cost subsides; 1.10 v.s. 0.49 for conditional migration cost subsidies).

Table 2.15 shows the moments of earnings and choices before and after introducing the government policies. The effects of policies on rural migrants' earnings are different. The lump sum subsidy for migration costs decreases the average migration costs. The individuals who are constrained by high migration costs are more likely affected by this policy. The average earnings decrease since most new migrants have lower reservation earnings. Since the policy of unemployment benefit increase migrants' reservation earnings, as we expected, the average logarithm earnings are much higher than the average logarithm earnings under the other two policies.

The unemployment benefit policy increases the value of the unemployment state and therefore increases their reservation earnings. As a result, the average earnings for rural migrants are higher compared to the average urban earnings under the other two policies. Also under this policy, the fraction unemployed increases to 49% in urban areas and the job search duration increases to 3.6 months. Almost half of rural migrants are unemployed causing the government to have to pay much more to achieve the urbanization goal.

The column of unconditional migration costs gives the individuals' choices under the policy of subsidy for migration costs. First, the average migration costs are significantly reduced by the government lump sum subsidy. After introducing this policy, average migration costs are only 37.4% of those before this policy. Second, the policy of providing lump sum subsidy for migration costs does not generate the large return and repeated migration behaviours. Before this policy, the fraction of the rural individuals who migrate within the observed periods is 1.41% and this fraction increases to 2.87% after introducing this policy. One of the reasons

[30]Yuan is Chinese currency, which is equal to about 1/8 U.S. Dollar in 2000.

to explain this phenomena is that there is no subsidy for return migration costs. Another reason is that the value of staying in urban areas is large. The fraction of employed in urban areas shows 78% of rural migrants have jobs which also indicates why rural migrants do not have many return and repeated circular migration.

The fifth column gives the simulation results if the government provides the conditional lump sum subsidy for migration costs. Under this policy only the migrants with social networks in urban areas can get the subsidy. This policy encourages rural individuals to invest their social networks. Average migration costs are 35.5% of those without this policy which are even lower than the average migration costs after introducing the unconditional migration cost subsidy (i.e. 37.4%). Since more rural migrants have social networks, the job search duration is shorter and the fraction of employed is larger than the policy of providing an unconditional migration cost subsidy.

2.6 Conclusion

This paper explores how social networks affect individuals' migration decisions and subsequent labor market outcomes. I construct and structurally estimate a dynamic model of migration choices allowing for return and repeated circular migration. In any given period, individuals make location, employment, and social network investment choices. In the model, individuals can accumulate their human capital through location-specific work experience.

In order to distinguish the effects of social networks through two different channels, I allow for the presence of social network to have a direct effect on migration costs, as well as an indirect effect on labor outcomes via an effect on the job arrival rate. In the model, individuals can invest in their social networks to increase the probability of creating or sustaining social networks.

I use the Chinese Household Income Projects (2007-2009) panel data and estimate the model by maximum likelihood. The estimation results show that social networks affect individuals' migration choices and subsequent labor market outcomes through both channels: reducing migration costs and search frictions. Social networks reduce about 10% of migration costs and almost double the job arrival rate for rural migrants.

The decomposition exercises show that social networks affect individuals' migration behaviour more through the channel of reducing search frictions than through the channel of lowering migration costs. For example, if social networks only reduce migration costs, 16% of individuals will migrate; if social networks only increase the job arrival rate, 26% of individuals will migrate.

The decomposition results also display that individuals effectively use and invest in social networks to optimize their migration decisions and labor market outcomes. For example, most individuals who invest in their social networks are individuals in rural areas and unemployed individuals in urban areas.

Next, I also propose three different types of policies all with the goal of meeting a 60% urbanization rate by 2020. The policy simulations show that a migration cost subsidy policy will cost less than a policy of providing unemployment benefits in urban areas. When comparing the two models (with and without social network investment decisions), I find that if individuals are allowed to invest in their social networks, they can effectively respond to the status of social networks and try to invest to increase or keep their social networks. In particular, individuals will invest less when the government tries to reduce migration costs or search frictions in urban areas. To offset the individuals' responses, the government has to spend more to encourage rural people to migrate to urban areas. These results show that it is important to consider the different roles of social networks when studying migration decisions and policies intended to affect migration levels.

Table 2.15: Policy Simulation Results: Choices

	Before Policy	After Policies		
	Benchmark	Unenployment Benefits	Migration Costs (Unconditional)	Migration Cost (Conditional)
Urban Log Earnings	7.1517	7.1561	7.0157	7.0461
Rural Log Earnings	5.0257	4.9413	4.9720	4.7675
Fraction of Migrants	0.2734	0.5009	0.5006	0.5005
with Networks	0.2078	0.4062	0.3771	0.3919
without Networks	0.0656	0.0948	0.1235	0.1085
Fraction of Moving Migrants	0.0141	0.0336	0.0287	0.0249
Fraction of Employees in Urban Areas	0.8197	0.5137	0.7815	0.7902
Fraction of Unemployed in Urban Areas	0.1803	0.4863	0.2185	0.2098
Job Search Duration (Month)	2.1059	3.6260	2.0388	1.9990
Average Age of Migrants	34.1179	35.3713	33.3058	35.1877
Average Job Arrival Rate	0.1287	0.1307	0.1269	0.1300
Average Migration Costs	100.00%	107.05%	37.38%	35.54%

1. The column of benchmark shows the values simulated by the model without proposing the government policies.

2. This table provides the three different policies' simulation results, all of which achieve the same goal of urbanization rate (i.e. 60% by 2020)

3. The column 3-5 give the simulation results when introducing the monthly unemployment benefits, an unconditional lump sum subsidy for migration costs, and a conditional lump sum subsidy for migration costs.

4. Earnings are the mean log earnings.

5. Migrants include people who were born in rural areas and resided in urban cities who can be employed or unemployed.

6. I normalize the average migration costs for rural migrants in the model to 1 unit. Other average migration costs give the relative values compared to the unit. (i.e. 37.38% means that the average migration costs after introducing the lump sum subsidy for migration costs is 37.38% of the unit.)

Bibliography

Beringhaus, S. and H.G. Seifer-Vogt. 1988. "Temporary versus Permanent Migration: a Decision Theoretical Approach." *Journal of Population Economics* 1:195–211.

Bishop, Kelly. 2012. "A Dynamic Model of Location Choice and Hedonic Valuation." *Working paper* .

Blejer, M.I. and I. Goldberg. 1980. "Return Migration-expectations versus Reality: a Case Study of Western Immigrants to Israel." *Research in Population Economics* 2.

Borjas, G. and B. Bratsberg. 1996. "Who leaves? The Outmigration of the Foreign-born." *Review of Economics and Statistics* 78:165–176.

Borjas, George. 2000. "The Economic Progress of Immigrants." *Issues in the Economics of Immigration* :15–50.

Borjas, G.J. 1999. *The Economic Analysis of Immigration.*, vol. 3A. Amsterdam: North-Holland.

Brandt, Loren and Carsten Holz. 2006. "Spatial Price Differences in China: Estimates and Implications." *Economic Development and Cultural Change* 55:43–86.

Buchinsky, M, C Gotlibovski, and O Lifshitz. 2012. "Residential Location, Work Location, and Labor Market Outcomes of Immigrants in Israel." *forthcoming Economatrica* .

Burdett, Kiefer N.M. Mortensen D.T., K. and G.R. Neumann. 1984. "Earnings, unemployment, and the allocation of time over time." *Review of Economic Studies* 51(4):559–78.

Carrington, Detragiache, and Vishwanath. 1996. "Migration with Endogenous Moving Costs." *American Economic Review* Vol. 86(4):909–930.

Chan, Kam Wing. 2012. *Internal Labor Migration in China: Trends, Geography and Polices*. New York: United Nations.

Chiswick, Lee, and Miller. 2005. "Family matters: the role of the family in immigrants' destination language acquisition." *Journal of Population Economics* 18:631–647.

Chiswick, B.R. 1978. “The effect of Americanization on the Earning of Foreign-born Men.” *Journal of Political Economy* 86:897–921.

Colussi, Aldo. 2006. “Migrants’ Networks: An Estimable Model of Illegal Mexican Imiigration.” *Working Paper* .

Edin, Fredriksson, and Aslund. 2003. “Ethnic Enclaves And The Economic Success Of Immigrants-Evidence From A Natural Experiment.” *The Quarterly Journal of Economics* 118:329–357.

Flinn, C. and J. Heckman. 1982. *Models for the Analysis of Labour Force Dynamics*. Greenwich: JAI Press.

Frijters, P and Bas van der Klaauw. 2006. “Job Search with Nonparticiparion.” *The Economic Journal* 116:45–83.

Galor, O and Stark O. 1990. “Migrants’ Savings, the Probability of Return Migration, and Migrants.” *International Economic Review* 31:463–467.

Gemici, Ahu. 2011. “Family Migration and Labor Market Outcomes.” *Working paper* .

Goel, Deepti and Kevin Lang. 2012. “Social Ties and the Job Search of Recent Immigrants.” *Working Paper* .

Hare, Denise and Shukai Zhao. 2000. “Lobor Migration as a Rural Development Strategy: a View from the Migration Origin.” *Working paper* .

Hare, Densise. 1999. “’Push’ versus ’Pull’ Factors in Migration Outflows and Returns: Determinants of Migration Status and Spell Duration among China’s Rural Population.” *Journal of Development Studies* 35:45–72.

Harris and M. P. Todaro. 1970. “Migration, unemployment and development: A two sector analysis.” *American Economic Review* Vol. 60:126–142.

Heckman, James and B. Singer. 1984. “A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data.” *Econometrica* 52:271–320.

Hu, Zhijun, Chongming Liu, and Zhimin Gong. 2011. “Estimation of the Gini Coefficient in China: 1985-2008.” *China Economic Quarterly* 10:1423–1436.

Huang, P. and F.N. Pieke. 2003. “China Migration Country Study.”

Jasso, G. and M.R. Rosenzweig. 1982. “Estimating the Emigration Rates of Legal Immigrants Using Administrative and Survey Data: the 1971 Cohort of Immigrants to the US.” *Demography* 19:279–290.

Keane and K. Wolpin. 1994. “The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence.” *The Review of Economics and Statistics* Vol. 76 (4):648–672.

———. 1997. “The Career Decisions of Young Men.” *Journal of Political Economy* Vol. 105(3):473–522.

Kennan and Walker. 2011. “The Effect of Expected Income on Individual Migration Decisions.” *Econometrica* Vol. 79:211–251.

Kono. 2006. “Employment with connections: Negative network effects.” *Journal of Development Economics* Vol. 81 (1):244–258.

Lessem, Rebecca. 2011. “Labor Migration and Wage Growth in Malaysia.” *Working paper* .

Liang, Zai. 2006. *Internal Migration: Policy Changes, Recent Trends and New Challenges. In Transition and challenge: China’s population at the beginning of the 21st century.* New York: Oxford Unveristy Press.

Lin, F. Cai, Justin and Z. Li. 1996. *The China Miracle: Development Strategy and Economic Reform*. Hong Kong: Chinese University Press.

Machin, S. and A. Manning. 1999. *Handbook of Labor Economics*, vol. 3C, chap. The Cause and Consequences of Longterm Unemployment in Europe. Amsterdam: North-Holland, 3085–139.

McCall, B. 1996. “Unemployment Insurance Rules, Joblessness, and Part-time Work.” *Econometrica* 64:647–682.

Meng, Xin. 2000. *Labout Market Reform in China*. Cambridge University Press.

Montgomery, J.D. 1991. “Social Networks and Labor-Market Outcomes: Toward an Economic Analysis.” *American Economic Review* 81(5):1408–1418.

Mortensen, D.T. 1986. “Job Search and Labor Market Analysis.” *Handbook of Labor Economics* 2:849–919.

Mortensen, D.T. and G.R. Neuman. 1984. *Choice or Chance? A Structural Interpretation of Individual Labor Market Dynamics*. Berlin: Springer-Verlag, 98–131.

Munshi, K. 2011. “Labor and Credit Networks in Developing Economies.” *Handbook of Social Economics* 1B:1223–1254.

Munshi, Kaivan. 2003. “Networks in the Modern Economy: Mexican Migrants in the U.S. Labour Market.” *Quarterly Journal of Economics* 118(2):549–599.

Murphy, Rachel. 1999. “Return Migration Entrepreneurs and Economic Diversification in Two Counties in South Jiangxu, China.” *Journal of International Development* 11:662–672.

Thom, Kevin. 2010. "Repeated Circular Migration: Theory and Evidence from Undocumented Migrants." *Working paper* .

Todaro, M. P. 1969. "A Model of Labor Migration and Urban Unemployment in Less Development Countries." *American Economic Review* 59:138–48.

Uhlendorff, A. and K.F. Zimmermann. 2006. "Unemployment dynamics among Migrants and Natives." *Working paper* .

Van den Berg, G. 2001. "Duaration Models: Specification, identication, and multiple duration." *Handbook of Econometrics* 5:3381–460.

Van den Berg, G.J. 1990. "Search Behavior, Transitions to Nonparticiparion and the Duration of Unemployment." *Economic Journal* 100:842–65.

Weiner, S.E. 1984. *A Survival Analysis of Adult Male Black-White Labor Market Flows*. Berlin: Springer-Verlag, 132–57.

Wolpin, K.I. 1987. "Estimating a Structural Search Model: the Transition From School to Work." *Econometrica* 55(4):801–17.

Wu, Harry Xiaoying. 1994. "Rural to Urban Migration in the People's Republic of China." *China Quarterly* 139:669–98.

Yao, Y. 2001. "Egalitarian Land Distribution and Labor Migration in Rural China." *working paper* .

Zhang, Junfu and Zhong Zhao. 2011. "Social-Family Network and Self- Employment: Evidence from Temporary Rural-Urban Migrants in China." *working paper* .

Zhao, Yaohui. 1997. "Labor Migration and Returns to Rural Education in China." *American Journal of Agricultural Economics* 79:1278–87.

———. 2000. "Rural-to-Urban Labor Migration in China: the Past and the Present." .

———. 2002. "Causes and Consequences of Return Migration: Recent Evidence from China." *Journal of Comparative Economics* 30:376–394.

———. 2003. "The Role of Migrant Networks in Labour Migration: the Case of China." *Contemporary Economic Policy* 21(4):500–511.

Zhao, Zhong. 2005. "Migration, Labor Market Flexibility, and Wage Determination in China: A Review." *Developing Economics* XLIII-2:285–312.

Chapter 3

Education Choices, Information and Borrowing Constraint

3.1 Introduction

Schooling, particularly college, is considered one of the main sources of human capital for the individual. Consequently, understanding the determinants of schooling attendance is, and has been for a while, an active area of research. There is, however, little work analyzing the interaction between the various determinants: ability, uncertainty, preferences and credit constraints. Different branches of the literature focus on estimating returns to schooling (e.g.,Ashenfelter and Krueger, 1994; Heckman and Vytlacil, 1998; Card, 2001); ability and returns to education (e.g., Cawley, Heckman, Lochner, and Vytlacil, 2000; Taber, 2001; Belzil and Hansen, 2002); the importance of parental income and, more generally, of borrowing constraints (e.g., Kane, 1996; Carneiro and Heckman, 2002; Cameron and Taber, 2004; Brown, Scholz, and Seshadri, 2012; Lochner and Alexander, 2011), etc. In most cases, the literature ignores the role played by the uncertainty facing the agent and uses ex-post measures (e.g. earnings at age 40) to analyze the agent's schooling decision. When they account for uncertainty, they assume that the unobserved (to the analyst) variability and the uncertainty facing the agent essentially coincide.[1]

In this paper, we contribute to the literature modeling multiple determinants of the schooling decision (e.g., Cameron and Heckman, 1998; Keane and Wolpin, 2001; Cameron and Heckman, 2001). We use economic theory and estimates of a semiparametrically identified structural model to analyze the role played by un-

[1]For example, both Keane and Wolpin (2001), Cameron and Heckman (2001) estimate their dynamic models of schooling assuming what is known by the agent and what is known by the econometrician at each point in time. While they allow for unobserved heterogeneity, it is essentially treated as an initial condition. Once the econometrician conditions on the initial heterogeneity, the evolution of the information set of the agent is given. However, there is no prior reason why one should assume that the evolution of what is unknown to the analyst and what is unknown to the agent coincide.

certainty and its interaction with ability, credit constraints and preferences in explaining college graduation. We estimate a structural model of schooling choice, labor supply, and consumption allocation under uncertainty, in which borrowing constraints arise from repayment constrains using pooled data from NLSY79 and PSID on white males.

One of the main contributions of this paper is to adapt the insight of Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005) to develop and implement a methodology that distinguishes information unknown to the econometrician but forecastable by the agent and information unknown to both at each period. The key to measuring uncertainty is to notice that individual choices reflect all the information known to the agent at a given time.[2] Responses of current decisions to future outcome innovations can be used to infer how much information the agent has. The semiparametric nature of the proposed test allows it to be used independently of the particular specification of the model as long as one considers families of models with the same determinants of choice.

The results we obtain do not single out any one particular determinant as the main reason why some people go to college and some don't. That this is the case, i.e., that all aspects of the problem play an important role, should not be surprising given the nature of the decision.

The key empirical results in the paper are:

1. At the time schooling decisions are made, wages are predictable. In particular, the estimates of the model imply that 52% of the unexplained variance in college log wages is predictable by the agent at age 18. This fraction is 56% for high school. This is similar to the results obtained by Cunha, Heckman, and Navarro (2005), Guvenen (2007). In fact, while the total unobserved variance of college log wages is higher than that of high school (i.e the unobserved variance from the analyst's perspective), the variance of the uncertain components of wages becomes proportionately even larger for college compared to high school under our estimated information set for the agent.

2. Once credit constraints are properly defined and relaxed, they play a more important role than previously estimated in the literature. When people are allowed to smooth consumption perfectly, i.e., when both credit constraints and uncertainty are completely eliminated, college attendance increases by 9.6%-points. We decompose this effect into the pure uncertainty effect, which accounts for almost two thirds of the increase, while the rest is due just to credit constraints. This result is not inconsistent with the evidence presented in Cameron and Heckman (2001), Keane and Wolpin (2001), Carneiro and Heckman (2002) and Cameron and Taber (2004) where credit constraints are found to be relatively unimportant. The analysis in these papers focuses on

[2] A similar idea motivates the work on the permanent income hypothesis of Flavin (1981) and Pistaferri (2001).

the inability of individuals to obtain funds to pay for tuition as the credit constraint. The relatively large effect of credit constraints may be in part due to the partial equilibrium nature of the exercise.

3. Agents also have preferences over schooling beyond the consumption value of earnings which we capture via an additive "psychic" cost function. Ability is one of the main determinants of costs and, as such, plays a key role in determining schooling decisions. High ability individuals face very low costs, while low ability individuals face large costs of attending college. This gives rise to schooling sorting by ability.

4. Schooling decisions are made by agents before all the relevant information about future outcomes has been revealed. Individual choices are made in an environment of uncertainty and agents base their decisions on their expectations and not on the realized outcomes observed by the econometrician. Expectations and realizations need not coincide. In particular, we estimate that eliminating uncertainty entirely – but keeping the credit constraints in place, 21% of high school graduates would instead choose to be college graduates and 11% of college graduates would regret their choice under uncertainty and pick high school instead. As a result, aggregate college attendance rises from 48.3% under uncertainty, to 54%.

This paper contributes to the literature on schooling choice by explicitly looking at the role played by uncertainty as a determinant of schooling. The idea is closely related to the work of Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005) in which a similar methodology is applied to extract agent's information at the schooling date. While Carneiro, Hansen, and Heckman (2003) assume no credit markets operate, and Cunha, Heckman, and Navarro (2005) assume an economy with perfect credit markets; we investigate an intermediate economy in which some credit markets operate and borrowing constraints arise from repayment constraints.[3] This allows for a more general setting in which consumption is not equal to income every period, nor is it necessary to assume that complete markets operate.

Our work also contributes to the literature on credit constraints and education.[4] In our analysis, credit constraints arise as a consequence of repayment restrictions (i.e., people cannot die in debt) and uncertainty about individuals future income. Our agents make decisions on schooling, labor supply, and asset holdings. Keane and Wolpin (2001) also consider marriage, residency with parents and the like. Our simpler model is, hopefully, more easily interpretable and lets us focus on assumptions about the information structure and identification, topics they do not consider.[5] Keane and Wolpin (2001) and Gourinchas and Parker (2002) (among others) assume that shocks to outcomes are unobservable to the econometrician

[3]See Laitner (1992), Aiyagari (1994) and Gourinchas and Parker (2002).

[4]See Lochner and Alexander (2012), Cameron and Taber (2004), Keane and Wolpin (2001).

[5]They do allow for variability to be different from uncertainty by allowing for unobserved types.

and uncertain to the agent. In our analysis, we infer the amount of uncertainty facing the agent from individual choices.

We also contribute to the literature on the estimation of the Frisch elasticity of labor supply. By focusing on longer periods, we essentially eliminate the need to distinguish between extensive and intensive margins that has been the focus of the recent literature (see Peterman, 2016). In fact, our estimated Frisch elasticity of labor supply of 1 is in the middle of the estimates obtained from both the Macro and Micro literatures.

Finally, we also contribute to the literature on consumption inequality and partial insurance. By extending the insight developed by Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005) and applying it to our model, we can identify what constitutes uncertainty at any stage in the life cycle.[6] A similar idea lies behind the test of the permanent income hypothesis in Flavin (1981). She picks a particular assumed ARMA (p, q) time series process for income and tests whether transitory income predicts consumption. Blundell and Preston (1998) and Blundell, Pistaferri, and Preston (2004) use the same idea to test for "partial insurance". As they acknowledge, their estimate of partial insurance combines the effects of information known to the agent but unknown to the econometrician and insurance. The methodology proposed in this paper can, in principle, distinguish between these two explanations.

The rest of the paper proceeds as follows. Section 3.2 presents a general version of the model of consumption allocation and schooling decisions that we use in the rest of the paper. The methodology used to infer the elements of the agent's information set is explained in section 3.3. In passing, we briefly sketch how semiparametric identification of the model is achieved. Appendix 1 presents a formal identification analysis. Section 3.4 describes the data we use and the parametrization of the model. In section 3.5 we present our analysis of what is in the agent's information set, and the empirical results from estimating the model using the right information. Section 3.6 concludes.

3.2 The Model

3.2.1 The Decision Process

Let $W_{i,s,t}$ denote the wage at time t of an individual i who has schooling level $s \in \{hs, col\}$ (hs=high school, col=college), $A_{i,t}$ denote the assets he saves for the next

This however, does not evolve over time. Conditional on the unobserved type which is given at the beginning of time, variability=uncertainty.

[6] Pistaferri (2001) and Kaufmann and Pistaferri (2009) use a similar idea by looking at expected wages as measured through a survey and measured wages in a consumption analysis.

period, $n_{i,t} \in [0,1]$ denotes the fraction of the time the individual spends working, $H_{i,t} = \bar{H}n_{i,t}$ denote total hours worked (out of a max of $\bar{H}$ possible hours), $Y_{i,s,t} = W_{i,s,t}H_{i,t}$ denote the earnings of the individual i, $u(C_{i,t}, H_{i,t})$ denote individual utility if the agent consumes $C_{i,t}$ and works $H_{i,t}$, $\widetilde{u}(A_{i,T})$ denote the utility in the terminal period when an individual reaches the last period with asset level $A_{i,T}$, and ρ be the discount rate. Let $\mathcal{I}_{i,t}$ be the information available to the agent at time t,[7] which is assumed to include all the past and current realizations of wages, the (assumed constant) interest rate r, and the asset stock. It may also contain some information about future wages. Exactly how much is what we try to determine in this paper.

Individuals live for $T+1$ periods and maximize expected lifetime utility in a world in which all risks arise from labor market risk and are idiosyncratic. At period $t = 0$, for each fixed schooling level s and given the (expected) wage sequence associated with each s, agents select an optimal intertemporal consumption allocation rule, as well as a rule for the proportion of time spent working. Agents can save and borrow as much as they want, subject to repayment constraints, via a single riskless asset A that pays a return r. Agents then make the schooling choice that maximizes expected utility.

More precisely, given schooling level s, the agent's problem at period $t > 0$ is to select how much of his time to spend working, and, given available resources, how much to consume (which determines how much to transfer to the next period). The value function given information set $\mathcal{I}_{i,t}$ is

$$\mathcal{V}_{i,s,t}(\mathcal{I}_{i,t}) = max_{H_{i,t},A_{i,t}} u(C_{i,t}, H_{i,t}) + \frac{1}{1+\rho} E(\mathcal{V}_{i,s,t+1}(\mathcal{I}_{i,t+1}) \mid \mathcal{I}_{i,t}) \tag{3.1}$$

$$s.t.\ C_{i,t} = Y_{i,s,t} + (1+r)A_{i,t-1} - A_{i,t} - \mathbb{1}(t=1, s=col)D_{i,t}, \tag{3.2}$$

$$A_{i,T} \geq -Y^{MIN}_{i,s,T+1}/(1+r). \tag{3.3}$$

$\mathbb{1}(t=1, s=c)$ is an indicator function that takes value one if the individual is in school, and $D_{i,t}$ is the direct cost of schooling (tuition) for individual i.

If the utility function satisfies standard conditions (*i.e.*, $lim_{C\to 0}u'(C,H) = \infty$ and concavity), the restriction that the agent cannot die in debt $\left(A_{i,T} \geq -Y^{MIN}_{i,s,T+1}/(1+r)\right)$ imposes a borrowing constraint on the individual at every period. In this case, the minimum value that assets can take at any period t (*i.e.*, the maximum amount the agent can borrow) is

$$A^{MIN}_{i,t} = \frac{A^{MIN}_{i,t+1} - Y^{MIN}_{i,s,t+1}}{1+r}, \tag{3.4}$$

[7]By information we mean the minimum sigma algebra generated by the random variables in $\mathcal{I}_{i,t}$. Since we associate the information set with a particular group of random variables $\mathcal{I}_{i,t}$, we use these concepts interchangeably.

where $Y^{MIN}_{i,s,t}$ is the minimum certain value that income can take at time t.

From the agent's perspective (i.e., given his information at time t) the solution to the maximization problem in equation (3.1) consists of a trio of time-schooling indexed functions: a policy function that tells him how much to work this period

$$H^*_{i,t} = \mathscr{H}_{s,t}\left(\mathcal{I}_{i,t}\right), \tag{3.5}$$

a policy function that tells him how much to save this period

$$C^*_{i,t} = \mathscr{C}_{s,t}\left(\mathcal{I}_{i,t}\right), \tag{3.6}$$

and the value function

$$V^*_{i,t} = \mathscr{V}_{s,t}\left(\mathcal{I}_{i,t}\right) \tag{3.7}$$

that gives the utility of working $H^*_{i,t}$ and consuming $C^*_{i,t}$ this period, and then following his optimal rules for $\tau > t$.

Once the agent solves the labor/consumption allocation problem and gets the value associated with each s at time $t = 1$, he uses it to select a schooling level. At period $t = 0$, the agent selects the schooling level s that maximizes his expected utility net of "psychic" costs, P. He will attend college if

$$E\left(\mathscr{V}_{col,1}\left(\mathcal{I}_{i,1}\right) - \mathscr{V}_{hs,1}\left(\mathcal{I}_{i,1}\right) - P_i \mid \mathcal{I}_{i,0}\right) > 0. \tag{3.8}$$

3.2.2 Specification of the Model

Wages for individual i at time t at schooling level s are written as[8]

$$\ln W_{i,s,t} = \mu_{s,t}\left(X_{i,t}\right) + U_{i,s,t}, \tag{3.9}$$

where $X_{i,t}$ represents variables that the econometrician observes and $U_{i,s,t}$ variables he cannot observe. We assume that the agent knows all of the variables in X at all times and that $U_{i,s,t}$ is revealed to him at period t. He may also know all or part of each $(U_{i,s,\tau}, \tau = t+1, ..., T)$ at time t. Uncertainty is thus associated with $\{U_{i,s,\tau}\}^T_{\tau=t+1}$. $U_{i,s,t}$ may also include measurement error in earnings. If this is the case, our estimates of uncertainty will be an upper bound since a fraction of the variance in $U_{i,s,t}$ will be due to the measurement error.

We write psychic costs in the schooling choice equation (3.8) as a function of variables Z that are observed by both the analyst and the agent. ζ represents variables not observed by the econometrician and may be (partially) known to the agent at $t = 0$. The net cost of schooling is

$$P_i = \phi\left(Z_i\right) + \zeta_i. \tag{3.10}$$

[8]Although not pursued in this paper, the separability assumption is not essential and can be relaxed using the analysis of Matzkin (2003) to analyze functions of the form $\ln W_{i,s,t} = \mu_{s,t}\left(X_{i,t}, U_{i,s,t}\right)$.

The utility function is of the CRRA form

$$u(C,H) = \frac{C^{1-\psi}}{1-\psi} - h\frac{n^{1+\frac{1}{\varphi}}}{1+\frac{1}{\varphi}}, \tag{3.11}$$

where $n = {}^{H}/_{\bar{H}}$, $\psi \geq 0$ is the coefficient of relative risk aversion, $h > 0$ weights the utility of leisure, and $\varphi \geq 0$ is the Frisch elasticity of labor supply. Since we do not model retirement explicitly, we assume that in the terminal period (i.e., the period after age 65) the utility function is given by

$$\widetilde{u}(A_{i,T}) = b\frac{\left(Y^{MIN}_{i,s,T+1} + (1+r)A_{i,T}\right)^{1-\chi}}{1-\chi}, \tag{3.12}$$

where b weights the utility in the terminal period, and we let the function have a different curvature parameter χ.

Finally, we assume that the data on both hours and consumption contain measurement error. For the case of labor supply we have

$$\ln\widehat{H}_{i,t} = \ln H_{i,t} + \delta_H K^H_{i,t} + \xi^H_{i,t}, \tag{3.13}$$

while for the case of consumption we have

$$\ln\widehat{C}_{i,t} = \ln C_{i,t} + \delta_C K^C_{i,t} + \xi^C_{i,t}. \tag{3.14}$$

Here $H_{i,t}, C_{i,t}$ are "true" hours and consumption, $\widehat{H}_{i,t}, \widehat{C}_{i,t}$ are measured hours and consumption, and measurement error that depends on observable variables $K_{i,t}$ and unobservable variables $\xi_{i,t}$. We allow for measurement error to depend on observables as we combine different datasets, and, for the case of consumption it is a natural assumption in this context since consumptions is measured at the household level, and from there they are imputed to the individual.[9] $K_{i,t}$ includes variables that control for the household structure as well as dataset of origin. $\xi_{i,t}$ is an unobserved term assumed to capture the rest.

3.3 Inferring the Agent's Information Set

The econometrician must know $\mathcal{I}_{i,t}$ in order to solve the model and develop estimating equations. Any conclusion extracted from the model relies crucially on the assumptions made about what constitutes the uncertainty facing the agent. It is thus important to develop a procedure to allow the analyst to separate the components of the agent's information set from what is unknown to him. We now turn our attention to this topic, sketching identification of the model in the process. Appendix D provides formal proof of identification.

[9]Imputation is done by dividing total consumption over the square root of the total number of members in the household.

3.3.1 Testing for Information Misspecification

We cast the problem of determining agent information sets as a testing problem.[10] We develop a simple test of misspecification for a proposed information set that does not depend on the details of the particular model being used. In this section we deal with the test of whether a candidate information set is correctly specified in a general setting.[11]

For any arbitrarily proposed (by the analyst) information set, $\widetilde{\mathcal{I}}_{i,t}$, it follows that $\ln \widehat{H}_{i,t} = \ln \mathscr{H}_{s,t}\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_H K_{i,t}^H + \xi_{i,t}^H$, and similarly for consumption. That is, measured hours should equal those predicted by the model via the policy function, $\mathscr{H}_{s,t}$, plus measurement error. Notice, however, that this holds true for a whole class of models (i.e., entire families of policy functions) besides the ones arising from the particular model we proposed in Section 3.2.

For a pair of nonparametric functions $\left(\mathscr{G}_{s,t}^H, \mathscr{G}_{s,t}^C\right)$ of the proposed information set, it is true that

$$\ln \widehat{H}_{i,t} = \mathscr{G}_{s,t}^H\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_H K_{i,t}^H + \xi_{i,t}^H, \tag{3.15}$$

$$\ln \widehat{C}_{i,t} = \mathscr{G}_{s,t}^C\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_C K_{i,t}^C + \xi_{i,t}^C. \tag{3.16}$$

The prediction that hours and consumption will be a function of the state variables of the model (i.e., the information set) is independent, for example, of the particular form of the utility function or of the wage equations.[12] It is in this sense that we can work with a nonparametric function of $\widetilde{\mathcal{I}}_{i,t}$, with the benefit that the solution of the dynamic program does not need to be computed, and that the test will still be valid for a general class of models predicting that hours and consumption can be written as in equations (3.15) and (3.16).

The test is simple: we want to estimate the model (either solving the dynamic problem for $\mathscr{H}, \mathscr{C}$ in equations (3.5) and (3.6), or by using nonparametric functions - polynomials on the elements of $\widetilde{\mathcal{I}}_{i,t}$ for example) using a candidate information set $\widetilde{\mathcal{I}}_{i,t}$. Equations (3.15) and (3.16) are part of the contribution of individual i, who selects schooling level s, to the likelihood. Alternatively, they could be used to form moments to compare against the data in a GMM setting for example.

Let $\pi_{\tau,t}^H$ and $\pi_{\tau,t}^C$ be a set of auxiliary parameters. To define the proposed test,

[10] See Cunha, Heckman, and Navarro, 2005 where a version of this test for a perfect credit markets model of schooling choice is proposed.

[11] The particular implementation of the test used in this paper is shown in section 3.3.3.

[12] The idea that we do not necessarily need to solve the whole dynamic program to estimate some functions is not new. See Hotz and Miller (1993), Aguirregabiria and Mira (2002) and Carranza (2007) for recent examples.

instead of basing the likelihood on equations (3.15) and (3.16), we use

$$\ln \widehat{H}_{i,t} = \mathcal{G}^{H}_{s,t}\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_H K^{H}_{i,t} + \xi^{H}_{i,t} + \sum_{\tau=t+1}^{T}\left[Y_{i,s,\tau} - E\left(Y_{i,s,\tau}|\widetilde{\mathcal{I}}_{i,t}\right)\right]\pi^{H}_{\tau,t},$$
$$\ln \widehat{C}_{i,t} = \mathcal{G}^{C}_{s,t}\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_C K^{C}_{i,t} + \xi^{C}_{i,t} + \sum_{\tau=t+1}^{T}\left[Y_{i,s,\tau} - E\left(Y_{i,s,\tau}|\widetilde{\mathcal{I}}_{i,t}\right)\right]\pi^{C}_{\tau,t}.$$

By assumption, the predicted hours and consumption $\left(\mathcal{G}^{H}_{s,t}\left(\widetilde{\mathcal{I}}_{i,t}\right), \mathcal{G}^{C}_{s,t}\left(\widetilde{\mathcal{I}}_{i,t}\right)\right)$ will not depend on the earnings innovations included in the last term on the right hand side of each equation since the agent integrates them out. The actual decisions, however, will be a function of the true agent's information set at t which may contain elements of $\left\{Y_{i,s,\tau} - E\left(Y_{i,s,\tau}|\widetilde{\mathcal{I}}_{i,t}\right)\right\}_{\tau=t+1}^{T}$. A test of which of the auxiliary parameters multiplying the earnings innovations $\left\{\pi^{H}_{\tau,t}, \pi^{C}_{\tau,t}\right\}_{\tau=t+1}^{T}$ equal zero is then a test of whether the proposed agent's information set at time t is correctly specified.[13] Notice that, as done in Cunha, Heckman, and Navarro, 2005, the schooling choice (or any other choice) may also be used when testing.

Given the many ways one can propose information sets $\widetilde{\mathcal{I}}_{i,t}$, specially the elements of the set unobserved to the econometrician, the test may to be of limited practical value. The next sections present assumptions to make the test operational. The main intuition, that under a correct specification of the information set the information innovations should not predict current choices, remains regardless of the implementation.

3.3.2 The Factor Structure and the Arrival of Information

In order to separate unobserved (to the econometrician) variability from the uncertainty facing agents, it is useful to assume that the unobservables for agent i can be factor analyzed in the following way:

$$\begin{aligned} U_{i,s,t} &= \theta_i \alpha_{s,t} + \varepsilon_{i,s,t} \\ \zeta_i &= \theta_i \lambda + \omega_i \end{aligned} \tag{3.17}$$

where θ_i is a vector of mean zero mutually independent "factors", $\varepsilon_{i,s,t}$ and ω_i are also mean zero random variables called "uniquenesses". Uniquenesses, factors and measurement errors, $\left(\xi^{H}_{i,t}, \xi^{C}_{i,t}\right)$, are all assumed mutually independent of each other for all schooling levels s and time periods t. The factor structure assumption is a natural starting point in analyses like the one in this paper. The wage equation can be interpreted as a pricing equation. Elements of the vector θ_i represent missing variables (i.e., variables unobserved by the econometrician) that affect outcomes and the factor loadings their prices.

[13] And it can in fact be considered as a form of Sims' test of causality (Sims (1972)).

The equations in (3.17) are only a statistical decomposition and, by themselves, are not informative about what is known to the agent at period t. We interpret elements of θ_i as permanent shocks that hit and influence earnings at different points in time. They provide a useful device for extracting components of uncertainty from future outcomes.

For identification purposes, we assume that the factor structure is such that wages in the first period are affected only by the first element of θ_i (so the loadings $\alpha_{s,t,l}$ for $l > 1$ and $t > 1$ would all be zero). Earnings in period 2 are affected by the first two elements of θ_i, the next period by the first three and so on. These elements of θ_i would then be revealed to the agent through their effect on wages . However, the agent might be able to forecast elements of θ_i that affect future wages but do not affect past and currently observed outcomes. Let $\underline{\theta_i(t)} = (\theta_{i,1}, ..., \theta_{i,t})$ denote those elements of θ_i that affect wages at or before t, and let $\overline{\theta}_i(t) = (\theta_{i,t+1}, ..., \theta_{i,T})$ denote those elements of θ_i that affect wages after t. We further separate $\overline{\theta}_i(t)$ into two components, $\left(\bar{\theta}_i^k(t), \bar{\theta}_i^u(t)\right)$ where $\bar{\theta}_i^k(t)$ is known by the agent at time t so it is in $\mathcal{I}_{i,t}$ and $\bar{\theta}_i^u(t)$ is unknown by the agent at time t so it is not in $\mathcal{I}_{i,t}$.

The following assumptions are made about the arrival of information

(I-1) The information revelation process of the agent is such that he either knows element l of θ_i, $\theta_{i,l}$ $(l = 1, \dots L)$, or he does not. Revelation of information when it happens, is instantaneous.

(I-2) At period t, the agent observes his outcomes for the period and so he knows $\{\varepsilon_{i,s,\tau}\}_{\tau=1}^{t}$ and the elements of $\theta_i(t)$, that is those elements of θ_i that affect outcomes in that period (or in any previous periods). If $\theta_{i,l}$ affects outcomes at $\tau \leq t$, then it is known by the agent at time t.

(I-3) Agents have rational expectations so that the expectations they take and the mathematical expectation operator with respect to the actual distributions in the model coincide.

The rest of the information structure of the model is assumed to be such that the agent has knowledge of the parameters of the model (e.g., $\rho, \psi, \mu(X)$) as well as knowledge of the observables X_i, K_i, Z_i,[14] and the uniqueness in the cost function ω_i. The econometrician never observes θ_i. By assumption, $\{\varepsilon_{i,s,\tau}\}_{\tau=t+1}^{T}$ is not part of the agent's information set $\mathcal{I}_{i,t}$.

[14]This assumption can be relaxed by modeling the stochastic process that generates these variables. For example, Keane and Wolpin (1997) have one X, experience in each sector, the accumulation of which they model.

3.3.3 Determining the Information Set Under the Factor Structure

In this section we cast the problem of determining agent information sets of section (3.3.1) in terms of the factor structure. Using the assumptions just made,[15] We redefine the test of section (3.3.1) to test whether a candidate information set defined in terms of of $\bar{\theta}_i^k(t)$ and $\bar{\theta}_i^u(t)$ is correctly specified.

The test proposed in Section 3.3.1 consists of including the income innovations relative to a proposed information set $\widetilde{\mathcal{I}}_{i,t}$ and test whether they affect current choices. Given the factor structure and information assumptions **(I-1) - (I-3)**, the only source of wages innovations "knowable" to the agent is given by $\bar{\theta}_i(t)$ (i.e., the factors that affect wages only in periods after t). Armed with this intuition it is possible then to design a simplified version of the test that does not require re-specifying the information set many times.

Start by assuming that the agents do not know the different elements of θ_i until they learn them when they hit wages. That is, the model is estimated using the candidate information set $\widetilde{\mathcal{I}}_{i,t}$ that contains no elements of $\bar{\theta}_i(t)$ before time t. To define the proposed test, instead of basing the likelihood on equations (3.15) and (3.16) we use

$$
\begin{aligned}
\ln \widehat{H}_{i,t} &= \mathcal{G}_{s,t}^{H}\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_H K_{i,t}^{H} + \xi_{i,t}^{H} + \bar{\theta}_i(t)\,\pi_t^{H}, \\
\ln \widehat{C}_{i,t} &= \mathcal{G}_{s,t}^{C}\left(\widetilde{\mathcal{I}}_{i,t}\right) + \delta_C K_{i,t}^{C} + \xi_{i,t}^{C} + \bar{\theta}_i(t)\,\pi_t^{C}.
\end{aligned}
$$

By assumption, the predicted hours and asset $\left(\mathcal{G}_{s,t}^{H}\left(\widetilde{\mathcal{I}}_{i,t}\right), \mathcal{G}_{s,t}^{C}\left(\widetilde{\mathcal{I}}_{i,t}\right)\right)$ will not depend on $\bar{\theta}_i(t)$ since the agent integrates them out. The actual decisions, however, will be a function of the true agent's information set at t which may contain elements of $\bar{\theta}_i(t)$. In this case, the different elements of $\bar{\theta}_i(t)$.

A test of whether those elements of $\left(\pi_t^H, \pi_t^C\right)$ associated with coordinate l of $\bar{\theta}_i(t)$ equal zero is then a test of whether element l of $\bar{\theta}_i(t)$ belongs in the agent's information set at time t. That is, if the l^{th} element of $\bar{\theta}_i(t)$ is actually part of the agent's information set *it will* affect the hours and consumption decisions, and, as a consequence, the elements of π_t^C associated with it will be estimated to be different from zero. Since we assume that the agent's do not know $\bar{\theta}_i(t)$ at t, the first time element l appears to affect choices is when it becomes known to the agent. For example, if factor 4 shows up as affecting consumption decisions at time 2 (i.e., its associated π_2^C is estimated to be different from zero) we conclude that the agent knows factor 4 at time 2, even though it does not affect wages until 2 periods later.

There is nothing special about the hours and consumption decisions. Any other decision variable that depends on future outcomes can also be used. The basic idea

[15] The assumption that θ is independent of X, Z is also imposed.

is that if element l of $\overline{\theta}_i(t)$ belongs in the information set of the agent at time t and the agent acts on it, it will affect the choices he makes at t. In particular it will affect the hours and consumption decisions and we can test for it. The same idea can be applied to agent schooling choices. Hence, estimating the schooling choice part of the model under $\widetilde{\mathcal{I}}_{i,0}$ (with no elements of θ_i contained in it) and using[16]

$$E\left(\mathscr{V}_{col,1}\left(\mathcal{I}_{i,1}\right)-\mathscr{V}_{hs,1}\left(\mathcal{I}_{i,1}\right)-P_i \mid \mathcal{I}_{i,0}\right)+\theta_i\pi_0>0$$

as the college choice rule allows us to test for $\pi_0=0$ as a test for misspecification.

3.3.4 Identification of the Model

Formal semiparametric identification analysis of the factor model of equations (3.1) - (3.17) is established in Appendix 1. This section provides an intuitive sketch of the identification arguments used in the Appendix for the factor structure of log wages,[17] as well as establishing identification of the preference parameters. In this paper identification theory is used to understand what in principle can be recovered nonparametrically from the data. We then use flexible parametric forms to obtain estimates of our high dimensional econometric model. The question of identifiability is a separate issue and should be judged independently of the choice of parametric forms for estimation purposes.[18]

Wages are identified by adapting a version of the arguments in Carneiro, Hansen, and Heckman (2003) and Cooley Fruehwirth, Navarro, and Takahashi (2016) which we now sketch for the case in which θ_i is a scalar. We assume that the problem of selection (i.e., that we only observe college wages for college graduates and high school wages for high school graduates) is solved using the arguments in Appendix 1 which involve using variation in the Z_i to achieve limit sets. Without loss of generality, take the system of log-wage equations for high school $\ln W_{i,hs,t}=\mu_{hs,t}\left(X_{i,hs,t}\right)+\theta_i\alpha_{hs,t}+\varepsilon_{i,hs,t},\ t\geq 1$.

First, notice that the factor θ_i has no natural scale, *i.e.,* $\theta_i\alpha=\kappa\theta_i\frac{\alpha}{\kappa}$ for any constant κ, so it needs to be set by a normalization as does the sign of the factor loading. Normalizing one loading takes care of both problems. Suppose that we normalize the loading in the first period so that $\alpha_{hs,1}=1$. Next, assuming that X is independent of the error terms $\{U_{i,hs,t}\}_{t=1}^{T}$, form cross moments of high school log wages from the data over time. Solving the system of equations that comes from equating the data (left hand side) to the theoretical moments predicted by the factor structure, we

[16]Notice that, same as before, we can use a nonparametric function of $\widetilde{\mathcal{I}}_{i,1}$ instead of the actual solution to the dynamic problem.

[17]Identification of the parameters of the "psychic" cost function is also established in Appendix 1. See also Heckman and Navarro (2007).

[18]See Roehrig (1988), Heckman (2005) for more on this distinction

obtain the factor loadings on the factor from

$$\frac{E\left((\ln W_{i,hs,t} - \mu_{h,t}(X_{i,h,t}))(\ln W_{i,hs,t'} - \mu_{hs,t'}(X_{i,hs,t'}))^{k'} \mid X\right)}{E\left((\ln W_{i,hs,1} - \mu_{h,1}(X_{i,h,1}))(\ln W_{i,hs,t'} - \mu_{hs,t'}(X_{i,hs,t'}))^{k'} \mid X\right)} = \frac{\alpha_{hs,t}\alpha^{k}_{hs,t'}E\left(\theta^{k+1}\right)}{\alpha^{k}_{hs,t'}E\left(\theta^{k+1}\right)} = \alpha_{hs,t}, \qquad t \neq t'.$$

Given the loadings for all time periods, using Theorem D.0.2 in Appendix 1, the distributions of both θ_i and $\{\varepsilon_{i,hs,t}\}_{t=1}^{T}$ can be nonparametrically identified. Notice that, while we can form moments for high school wages over time, we can never form moments of log wages across schooling levels since wages are not observed on both schooling levels for anyone.

Given the normalizations we just made to the high school system of wages, making a similar set of normalizations to the college system would amount to setting the sign (and magnitude) of the unobserved covariance between college and high school log wages. To see this, notice that the unobserved covariance of log wages in high school and college in period 1 is $\mathrm{cov}(\ln W_{i,hs,1}, \ln W_{i,col,1} \mid X) = \alpha_{c,1}\sigma^2_{\theta}$, so setting $\alpha_{col,1} = 1$ would impose a strong restriction that the covariance is positive and fixed by the variance determined in the high school system. Theorem D.0.3 and Corollary D.0.4 in Appendix 1 show that restrictions of this nature do not need to be imposed.

Preferences, the discount factor, and measurement error in hours and consumption can be identified from the usual first order condition and Euler equation arguments. From the first order condition for labor supply we have

$$C^{-\psi}_{i,t} w_{i,t} \mathcal{H}^{1+\frac{1}{\varphi}} = hH^{\frac{1}{\varphi}}_{i,t}.$$

Replacing for consumption and labor by their measured (with error) counterparts and taking logs we have

$$\ln \widehat{H}_{i,t} = \varphi\left[\left(1+\frac{1}{\varphi}\right)\ln\mathcal{H} - \ln h\right] - \psi\varphi\ln\widehat{C}_{i,t} + \varphi\ln w_{i,t} + \delta_H K^H_{i,t} + \psi\varphi\delta_C K^C_{i,t} + \left(\xi^H_{i,t} + \psi\varphi\xi^C_{i,t}\right). \quad (3.18)$$

$\ln\widehat{C}_{i,t}$ in equation (3.18) is correlated with the residual via $\xi^C_{i,t}$. Under the assumption that measurement error is uncorrelated we can instrument for $\ln\widehat{C}_{i,t}$ with lagged $\ln\widehat{C}_{i,t-\tau}$ for $\tau > 0$. With an instrument in hand, equation (3.18) identifies φ, ψ, h, the unique elements of δ_H and δ_C, $(\delta_H + \psi\varphi\delta_C)$ for the common elements of $K_{i,t}$, as well as the convolution $\xi^H_{i,t} + \psi\varphi\xi^C_{i,t}$.

While the risk aversion parameter ψ is identified from the argument above, it can also be identified (along with the discount factors and the measurement error in consumption) using relatively standard Euler equation arguments.[19] Since the arguments are well known, what follows is a simple sketch of how this is done

[19] See Hansen and Singleton (1983), Browning and Lusardi (1996) and Attanasio and Low (2004).

assuming that $K_{i,t}^{C}, K_{i,t+1}^{C}$ are contained in $\mathcal{I}_{i,t}$. Using the first order condition of equation (3.7) and using equations (3.11) and (3.14) it follows that

$$E\left(\frac{1+r}{1+\rho}\left(\frac{\widehat{C}_{i,t+1}e^{K_{i,t}^{C}\delta_C+\xi_{i,t}^{C}}}{\widehat{C}_{i,t}e^{K_{i,t+1}^{C}\delta_C+\xi_{i,t+1}^{C}}}\right)^{-\psi}-1\mid\mathcal{I}_{i,t}\right)=0. \tag{3.19}$$

which is a standard consumption Euler equation except that it contains measurement error and so the standard argument of Hansen and Singleton (1983) cannot be applied directly. Instead, as noted by Chioda (2004),[20] one can take differences of two adjacent Euler equations to form a valid moment condition and identify ψ. With ψ in hand rewrite equation (3.19) as

$$\frac{1+r}{1+\rho}\left(\frac{\widehat{C}_{i,t+1}e^{K_{i,t}^{C}\delta_C+\xi_{i,t}^{C}}}{\widehat{C}_{i,t}e^{K_{i,t+1}^{C}\delta_C+\xi_{i,t+1}^{C}}}\right)^{-\psi}=\eta_{i,t}+1$$

where $\eta_{i,t}$ is expectational error which is a function (among other things) of the elements of θ_i not contained in $\mathcal{I}_{i,t}$. Taking logs and a linear approximation of $\eta_{i,t}+1$ around $\eta_{i,t}=0$ we obtain

$$\ln\frac{\widehat{C}_{i,t+1}}{\widehat{C}_{i,t}}=\frac{1}{\psi}\ln\left(\frac{1+r}{1+\rho}\right)+\left(K_{i,t+1}^{C}-K_{i,t}^{C}\right)\delta_C+\left(\xi_{i,t+1}^{C}-\xi_{i,t}^{C}-\frac{\eta_{i,t}}{\psi}\right) \tag{3.20}$$

From the fact that the interest rate r is given, it follows that we can identify the discount factor ρ and the observable effect of measurement error δ_C for those elements of $K_{i,t}^{C}$ that change over time.

We next proceed to look at the Euler equation in the terminal period

$$-\left(\widehat{C}_{i,T}e^{-\delta_C K_{i,t}^{C}-\xi_{i,t}^{C}}\right)^{-\psi}+\frac{b(1+r)}{1+\rho}\left(\underline{E}_{T+1}+(1+r)A_{i,T}\right)^{-\chi}=0 \tag{3.21}$$

where there is no expectation with respect to A_T since it is known at time T. For identification purposes it is helpful to rewrite equation (3.21) as

$$\ln\widehat{C}_{i,T}=-\frac{1}{\psi}\ln\left(\frac{b(1+r)}{1+\rho}\right)+\frac{\chi}{\psi}\ln\left(\underline{E}_{T+1}+(1+r)A_{i,T}\right)-K_{i,T}^{C}\delta_C-\xi_{i,T}^{C}. \tag{3.22}$$

Since ψ, and ρ are known from the argument above, b, χ, the remaining elements of δ_C as well as the distribution of $\xi_{i,T}^{C}$ are identified.

Once identification of the preference parameters is secured, all the elements required to solve the consumption allocation and the labor supply problem of equation (3.1) are in place. To show that the distributions of the unobserved part of measurement error are identified remember that the assumption that $\xi_{i,t}^{H}, \xi_{i,t}^{C}$ are independent of θ_i and of $K_{i,t}$ is imposed. Since the left hand side of $\ln\widehat{C}_{i,t}-K_{i,t}^{C}\delta_C=\ln C_t(\mathcal{I}_{i,t})+\xi_{i,t}^{C}$ is known and so is the distribution of $\ln C_t(\mathcal{I}_{i,t})$ we can recover the distribution of $\xi_{i,t}^{C}$ by deconvolution. By either following a similar argument, or noticing that $\xi_{i,t}^{H}+\psi\varphi\xi_{i,t}^{C}$ is identified from equation (3.18), it is easy to show that $\xi_{i,t}^{H}$ is identified.

[20] See also Ventura (1994) for the parametric case.

3.4 Data and Parametrization of the Model

The model in this paper is estimated on a sample of white males who either graduated high school (and only high school) or are college graduates. Since the PSID data does not contain enough observations with measurements for ability, and the NLSY79 does not contain the full lifecycle for any individual, the sample used contains individuals from both the NLSY79 and PSID datasets pooled together. The sample consists of a total of 1,642 white males born between 1923 and 1964, who either took the ASVAB battery of tests (NLSY) or the IQ Word Test (PSID). Of these, 1,263 come from NLSY79 and 379 from PSID.

Individual working life cycles are simplified to eight 6-year-long periods.[21] This simplifying assumption is used to keep the computational complexity of the model manageable.[22] For each period, hours worked is simply the sum of total hours worked during the six years. Log-earnings in the period are calculated as the log of the present value of earnings for the period discounted at $r = 3\%$. Wages for the period are then calculated as the ratio of earnings to hours worked in the period. Earnings and hours for individuals who are missing are imputed only in the years in which there was no survey using an average of the earnings in the years immediately adjacent (i.e., the year before and the year after) to the missing year. If earnings are not available in either of these years they are left as missing. Missings are treated as random events.[23]

Table 3.1 presents a summary of the pooled dataset used to estimate the model.[24] In general, college graduates get higher wages, consume more and have more assets than high school graduates. They also have higher test scores, come from better family backgrounds, have fewer siblings and are more likely to live in a location where college tuition is lower.

For each schooling level $s = \{hs, col\}$ and for each period of wages $t = \{1, ..., 8\}$ we model $\ln W_{i,s,t}$ as being generated by a factor model:

[21]Period 1 covers ages 18 to 23, period 2 ages 24 to 29, period 3 ages 30 to 35, period 4 ages 36 to 41, period 5 ages 42 to 47, period 6 ages 48 to 53, period 7 ages 54 to 59 and period 8 ages 60 to 65. There is an additional terminal period in the model which includes whatever happens after age 65. Utility in this terminal period is modeled as depending only on assets carried to that period and a minimum income (e.g., social security income), so no additional data for the period is required.

[22]Estimation of the model requires the solution of the dynamic programming problem every time a different parameter vector is tried. Furthermore, the evaluation of the likelihood itself requires calculating a multidimensional integral (7 factors plus initial assets) for each individual in the sample.

[23]See Fitzgerald, Gottschalk, and Moffitt (1998) for evidence that observed people in PSID have similar characteristics as those in the CPS so attrition is roughly random. See MaCurdy, Mroz, and Gritz (1998) for evidence that attrition is roughly random in the NLSY79.

[24]A more complete description of the dataset is presented in Appendix D.

Table 3.1: Descriptive Statistics: White Male High School and College Graduates

	Full Sample					High School					College Graduates				
Variable Name	**Obs**	**Mean**	**Std Dev**	**Min**	**Max**	**Obs**	**Mean**	**Std Dev**	**Min**	**Max**	**Obs**	**Mean**	**Std Dev**	**Min**	**Max**
Log Wages	5,813	-6.44	0.64	-10.01	-3.70	2,893	-6.64	0.53	-9.86	-4.09	2,920	-6.24	0.68	-10.01	-3.70
PSID	13,136	0.23	0.42	0	1	6,768	0.21	0.40	0	1	6,368	0.26	0.44	0	1
Mother's education	13,136	4.22	1.60	0	8	6,768	3.73	1.29	0	8	6,368	4.73	1.72	0	8
Year of Birth	13,136	1955.73	10.16	1923	1964	6,768	1956.20	9.91	1923	1964	6,368	1955.24	10.40	1923	1964
Number of Siblings	13,136	2.85	2.00	0	18	6,768	3.22	2.17	0	18	6,368	2.45	1.71	0	12
PV of Local Tuition at age 18	13,136	0.72	0.27	0.00	2.01	6,768	0.74	0.28	0	1.80	6,368	0.69	0.26	0.15	2.01
Local Unemp at age 17: High School	13,136	0.06	0.03	0.00	0.54	6,768	0.06	0.03	0.00	0.54	6,368	0.06	0.03	0.00	0.22
Local Unemp at age 17: College	13,136	0.03	0.02	0.00	0.35	6,768	0.03	0.02	0.00	0.35	6,368	0.03	0.02	0.00	0.16
College	13,136	0.48	0.50	0.00	1.00	6,768	0	0	0	0	6,368	1	1	1	1
ASVAB															
Arithmetic Reasoning	10,104	0.53	0.30	0.03	1	5,376	0.38	0.25	0.03	1	4,728	0.70	0.26	0.03	1
Word Knowledge	10,104	0.53	0.31	0.03	1	5,376	0.38	0.27	0.03	1	4,728	0.71	0.25	0.03	1
Paragraph Composition	10,104	0.56	0.31	0.03	1	5,376	0.41	0.29	0.03	1	4,728	0.74	0.25	0.03	1
Coding Speed	10,104	0.51	0.29	0.03	1	5,376	0.41	0.27	0.03	0.98	4,728	0.63	0.27	0.03	1
Math Knowledge	10,104	0.53	0.30	0.03	1	5,376	0.35	0.22	0.03	1	4,728	0.74	0.24	0.03	1
IQ Word Test	3,032	0.58	0.29	0.01	1	1,392	0.48	0.28	0.01	1	1,640	0.67	0.27	0.03	1
Grade Completed at Test Date	13,136	12.07	2.40	7	21	6,768	11.08	1.27	7	14	6,368	13.11	2.83	8	21
Enrolled at Test Date	13,136	0.45	0.50	0	1	6,768	0.26	0.44	0	1	6,368	0.66	0.48	0	1
Age at Test Date	13,136	22.42	7.26	16	49	6,768	22.16	7.11	16	49	6,368	22.70	7.40	16	49
Ability	13,136	0.00	0.24	-0.82	0.52	6,768	-0.08	0.22	-0.52	0.52	6,368	0.08	0.23	-0.82	0.52
Consumption	2,835	25.71	29.80	0.18	365.85	1,416	19.14	21.55	0.18	358.57	1,419	32	35.00	0.45	365.85
Assets 18-65	5,903	12.52	29.79	-24.84	476.90	3,002	7.31	18.12	-24.84	427.40	2,901	17.90	37.55	-21.63	476.90
Hours	6,623	12517.57	4156.55	35.00	23952.00	3,330	12555.71	3948.28	40.00	23952.00	3,293	12479.00	4357.30	35.00	23940.00
Married	9,267	0.63	0.45	0	1	4,709	0.62	0.45	0	1	4,558	0.65	0.45	0	1
Number of Children	9,326	0.87	1.10	0	8.67	4,730	0.86	1.08	0	7	4,596	0.87	1.12	0.00	8.67

1. x 10,000. Measured in year 2000 US dollars.
2. Both ASVAB tests and IQ Word tests are measured in percentiles of their respective distributions.

$$\ln W_{i,s,t} = X_{i,t}\beta_{s,t} + \sum_{j=1}^{\min\{t,7\}} \theta_{i,j}\alpha_{s,t,j} + \varepsilon_{i,s,t}.$$

To pin down the scale of each θ_j, we normalize the loading of the high school wage equation to be one for that factor in the same period. That is, for factor $\ell = 1, ..., 7$ we normalize $\alpha_{h,\ell,\ell} = 1$. The psychic cost function is also allowed to (in principle, since only factors in the agent's information set at the schooling decision age affect it) be a function of all factors:

$$P_i = Z_i\gamma + \sum_{j=1}^{7} \theta_{i,j}\lambda_j + \omega_i.$$

The Z_i include variables that only affect the schooling decision like family background, and local unemployment.

In order to account for ability in a consistent way across both datasets, an additional one factor model that utilizes a system of external measurements on ability is used. For the case of NLSY79 data, five components from the ASVAB battery of tests, measured as percentile ranks in the population are used. For PSID, the 1972 IQ test (also measured as percentile ranks) is included.[25] Each test, $M_{i,j}$, is modeled as a function of individual ability, Q_i:

$$M_{i,j} = X_i^M\beta_j^M + Q_i\alpha_j^M + \varepsilon_{i,j}^M. \tag{3.23}$$

To pin down the scale and sign of ability, the loading on the arithmetic reasoning test $\left(\alpha_1^M\right)$ is normalized to 1. This normalization associates higher levels of the factor with higher test scores, purged of the effect of X^M, so we interpret Q_i as ability. In this interpretation, tests are assumed to be noisy proxies for ability which is given by Q_i. Identification of the loadings and non-parametric distributions of $Q_i, \varepsilon_{i,j}^M$ follows from the same arguments used in Theorem D.0.3. Table 3.2 shows the full set of covariates used for ability measures, log wages, costs, and measurement error in hours and consumption.

Each of the factors $\theta_{i,\ell}$ is allowed to follow a mixture of normals distribution.[26] In all cases mixtures with 2 elements are found to be adequate. The distribution of ability follows a mixture of three normals: $Q_i \sim \sum_{j=1}^{3} \pi_{Q,j} f\left(Q_i; \mu_{Q,j}, \sigma_{Q,j}^2\right)$. The remaining distributions in the model (e.g., measurement error, uniquenesses) are all assumed to be normal.

In order to allow for levels of borrowing that are roughly consistent with the ones observed in the data, we assume that individuals have access to some guaranteed income (e.g., social security income) when retired, $Y_{i,S,T+1}^{MIN}$. We set the levels

[25] See also Hansen, Heckman, and Mullen (2004) for an analysis of ASVAB tests and their relation to ability.

[26] $\theta_{i,\ell} \sim \sum_{j=1}^{J_\ell} \pi_{\ell,j} f\left(\theta_{i,\ell}; 0, \sigma_{\ell,j}^2\right)$, where $f\left(x; \mu, \sigma^2\right)$ is a normal density with mean μ and variance σ^2

Table 3.2: List of Covariates

	Wages	**Test System**	**Cost Function**	**Measurement Error**	
				Hours	Consumption
Age	Yes	No	No	Yes	Yes
Age squared	Yes	No	No	No	No
Ability	Yes	Yes	Yes	No	No
NLSY Dummy	Yes	No	Yes	Yes	Yes
Married	No	No	No	No	Yes
Number of Children	No	No	No	No	Yes
Age at Test Date	No	Yes	No	No	No
Grade Completed at Test Date	No	Yes	No	No	No
Enrolled at Test Date	No	Yes	No	No	No
Local Unemp at age 17: High School	No	No	Yes	No	No
Local Unemp at age 17: College	No	No	Yes	No	No
Number of Siblings	No	No	Yes	No	No
Mother's education	No	No	Yes	No	No

of $Y^{MIN}_{i,s,T+1}$ to be roughly consistent with the 18 additional years of life expectancy at age 65 observed in the data, as well as with the average levels of social security income retiree's got in 2016: 1,341 yearly for high school graduates (974 in 2000 dollars), and 2,000 yearly for college graduates (1,453 in 2000 dollars). With these numbers in place, in any given period, the maximum level of borrowing sustained in the model, $A^{MIN}_{i,s,t}$, is automatically determined by equation (3.4).

In our estimation, we set the maximum number of hours one can work,$\bar{H}$, to 24000 hours per-period for all periods except one. We only allow people who are in college in period 1 to work half time, and hence set $\bar{H}$=12000 in this case. We also approximate the (unobserved) distribution of initial assets (i.e., assets in period 0) with the distribution of assets in the first three periods multiplied times 0.5.

The model is estimated in two stages. In the first stage we estimate the model in equation (3.23) by maximum likelihood.[27] Given the estimates, we predict Q_i for all individuals and include it as part of the X and Z variables for the rest of the model.

Estimation of the rest of the model is done by maximum likelihood using a combination of accelerated random search, the Nelder-Meade simplex method and the BFGS algorithm to maximize the likelihood. The contribution of individual i who chooses schooling $S_i = s$

[27]The contribution to this likelihood for individual i is

$$\int_Q \prod_{j=1}^{J} f_{\varepsilon^M_{i,j}}\left(M_{i,j} - X^M_i \beta^M_j - Q_i \alpha^M_j | Q_i, X^M_{i,j}\right) dF(Q).$$

$$\int_{A_0,\Theta}\left[\begin{array}{c}\prod_{t=1}^{8} f_{\varepsilon_{i,s,t}}\left(\ln W_{i,s,t}-X_{i,t}\beta_{s,t}-\theta_i\alpha_{s,t}|\theta_i,X_{i,s,t}\right)\\ \prod_{t=1}^{8} f_{\xi^H_{i,t}}\left(\ln \widehat{H}_{i,t}-\ln \mathscr{H}_{s,t}\left(\mathcal{I}_{i,t}\right)-K^H_{i,t}\delta_H|\theta_i,X_{i,s,t},Z_i,K^H_{i,t}\right)\\ \prod_{t=2}^{8} f_{\xi^C_{i,t}}\left(\ln \widehat{C}_{i,t}-\ln \mathscr{C}_{s,t}\left(\mathcal{I}_{i,t}\right)-K^C_{i,t}\delta_C|\theta_i,X_{i,s,t},Z_i,K^C_{i,t}\right)\\ Pr\left(S_i=s|\theta_i,X_{i,s,t},Z_i\right)\end{array}\right]dF\left(A_0,\theta\right).$$

Evaluation of the likelihood requires that the econometrician solve the dynamic program (for a given proposed $\mathcal{I}_{i,t}$) in order to evaluate the schooling selection probability, the choice of hours worked, and the consumption policy function $\mathscr{C}_{s,t}$. Since the econometrician never observes any element of θ_i or assets in period 0, he has to integrate against their distribution when evaluating the likelihood. In the model we estimate, the value function given by the solution to the hours and consumption allocation problem is approximated numerically using a second order complete polynomial approximate.

3.5 Empirical Results

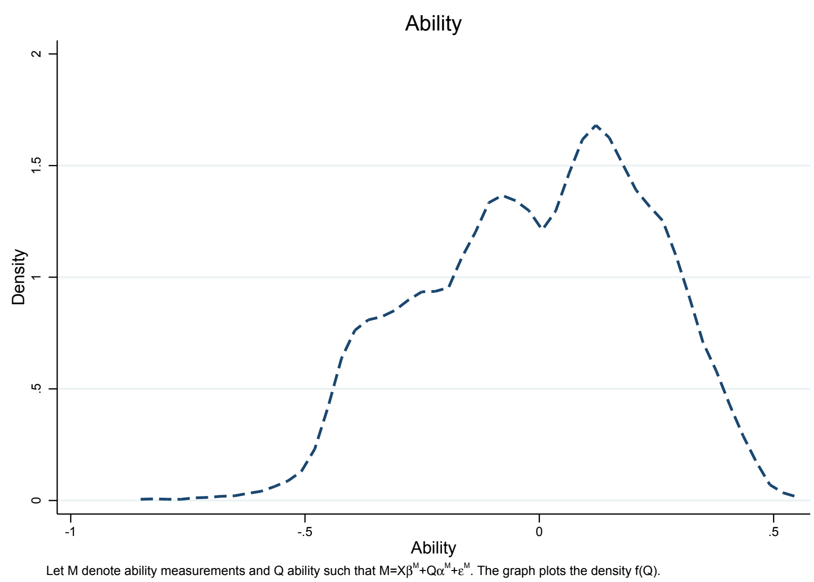


Figure 3.1: Distribution of Predicted Ability

3.5.1 Ability

We begin by estimating the model for ability of equation (3.23). We use five components of the ASVAB battery of tests contained in the NLSY-79: Arithmetic Reasoning, Word Knowledge, Paragraph Composition, Coding Speed and Math Knowledge; as well as the 1972 IQ Word Test for PSID. Tables A3-1 and A3-2 in Appendix

Table 3.3: Test for Information Set Misspecification

Auxiliary Parameters	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7
Schooling Choice	-0.970^+	-0.512^+	-0.095	-0.160	-0.126	0.064	0.388^+
	(0.091)	(0.144)	(0.120)	(0.146)	(0.105)	(0.180)	(0.154)
Log Hours							
Age 18-24		0.303^+	0.015	-0.007	0.065	0.269^+	-0.128^+
		(0.067)	(0.038)	(0.057)	(0.045)	(0.055)	(0.062)
Age 25-30			0.009	0.112^+	-0.027	0.132^+	-0.175^+
			(0.023)	(0.033)	(0.027)	(0.034)	(0.034)
Age 31-36				0.081^+	-0.038^+	0.102^+	-0.186^+
				(0.021)	(0.014)	(0.017)	(0.019)
Age 37-42					-0.044^+	0.110^+	-0.126^+
					(0.014)	(0.020)	(0.015)
Age 43-48						0.140^+	-0.101^+
						(0.037)	(0.031)
Age 49-54							-0.265^+
							(0.103)
Log Consumption							
Age 25-30			0.117	0.059	-0.096	-0.055	-0.103
			(0.064)	(0.097)	(0.066)	(0.080)	(0.078)
Age 31-36				0.130^+	-0.023	0.003	-0.277^+
				(0.061)	(0.036)	(0.049)	(0.051)
Age 37-42					-0.045	0.133	-0.048
					(0.061)	(0.074)	(0.068)
Age 43-48						0.075	-0.093
						(0.074)	(0.076)
Age 49-54							0.165
							(0.183)

1. Let $g(\mathcal{I})$ be the predicted choice as a function of the information set $\mathcal{I}$. The factor not included in $\mathcal{I}$, are added to the choice's contribution to the likelihood, and we test whether their associated parameters are different from zero.
2. Standard Errors in parenthesis.
3. + Significantly different from zero, at levels of 5% or less.

3 contain the parameter estimates for this model. As expected, ability is associated with higher test scores. With the estimates in hand, we then use Bayes rule to predict ability for each individual. Figure 3.1 plots the distribution of predicted ability that we obtain using this procedure. As can be seen from the graph, the distribution is highly non-normal, hence it is important to allow for more general distributions like the mixture distribution we use.

Table 3.4: Model Fit: Average Outcomes

	Log Hours		**Log Consumption**		**Log Wages**	
	Data	**Predicted**	**Data**	**Predicted**	**Data**	**Predicted**
Age 18-23	8.656	8.673			-7.307	-7.290
		[8.434, 8.788]				[-7.360, -7.244]
Age 24-29	9.395	9.408	2.825	2.819	-6.725	-6.762
		[9.286, 9.528]		[2.722, 2.923]		[-6.817, -6.720]
Age 30-36	9.512	9.473	2.857	2.860	-6.489	-6.502
		[9.418, 9.529]		[2.778, 2.945]		[-6.548, -6.458]
Age 36-41	9.546	9.519	2.898	2.870	-6.286	-6.323
		[9.458, 9.577]		[2.760, 2.982]		[-6.374, -6.277]
Age 42-47	9.493	9.521	3.102	3.107	-6.227	-6.243
		[9.439, 9.597]		[2.978, 3.222]		[-6.299, -6.193]
Age 48-53	9.456	9.450	3.338	3.339	-6.163	-6.200
		[9.260, 9.682]		[3.125, 3.630]		[-6.280, -6.130]
Age 54-59	9.372	9.352	3.267	3.248	-6.185	-6.218
		[9.146, 9.536]		[2.866, 3.608]		[-6.356, -6.098]
Age 60-65	8.863	8.872	3.436	3.434	-6.318	-6.348
		[8.589, 9.152]		[2.50, 4.128]		[-6.526, -6.194]

1. 95% Confidence interval in brackets.

3.5.2 Test of Misspecification

Table 3.3 presents the results of the proposed test of misspecification of the agent's information set using the auxiliary parameters, π, defined in Section 3.3.3 for the elements of $\bar{\theta}_i(t)$. We estimate all of the policy functions using a sieve for each function. In terms of the empirical model described above, this entails testing whether $\theta_{i,\ell}$ affects the agent's choices before period ℓ (when it hits wages).

In summary, the results of the test show that ability, θ_1, θ_2 and θ_7 are known at the time schooling choices are made; that θ_6 becomes known in period 1; additionally θ_4 becomes known in period 2; finally all factors are known at period 3. As a consequence, all of the results presented in the next sections are based on estimates of the structural model using this information set.

Tables A3-3 to A3-7 in Appendix E present the parameter estimates for the model estimated under this information structure. In total the model has 144 parameters.[28] The estimated parameters are all within reasonable ranges. In particular, we estimate a coefficient of risk aversion for consumption of 0.49 (on the low side), a Frisch elasticity of labor of 1.00 and a discount factor of 0.94, which implies a yearly discount rate of around 1%.

3.5.3 Model Fit to the Data

To validate the model estimates obtained under the information set chosen as a consequence of the test in the previous section, a variety of checks of fit of predictions of the model versus their data counterparts are performed. First, the proportion of people who attend college in the data and the one predicted by the model are compared.[29] Whereas 48.5% of the people in the sample are college graduates, the model predicts roughly 48.3% slightly below the actual number. The 95 confidence interval for this predicted proportion is $(44.7, 54.1)$, hence the null hypothesis of equality of predicted and actual proportions cannot be rejected.

Table 3.4 presents the (per-period) mean for the logs of wages, hours worked, and consumption both in the data as well as predicted from the model. Overall the model does a remarkably good job of matching the moments in the data. When we perform formal tests of equality between data and model means in Table 3.4, we cannot reject equality in any case. As we show in Tables A3-8 and A3-9 in Appendix 3, the model fits the data on consumption and wages by schooling level equally well, but has a much harder time reproducing the pattern of hours by schooling level.

As a final check on the predictions of the model we estimate a treatment-on-the-treated parameter using a semi-parametric control function approach as described

[28]The interest rate r is set at 3% annually.

[29]All of our predicted results are based on simulating (i.e., sampling the unobservables) 50 times for each individual in the data.

in Navarro (2008) on both the data and on data simulated from the model. We perform this comparison since the control function approach does not impose a factor structure on the errors, but the simulated data is generated under this assumption. In the first stage we predict the probability of going to college with a probit of schooling on local unemployment, tuition, number of siblings, mothers education, and ability for both the data and the simulated data. We then run a regression of log wages on age, age-squared, ability, an indicator for NLSY and a 2nd degree polynomial on the (log) probability of schooling, first for high school graduates and then for college graduates. These gives us unbiased estimate of the schooling specific wage regressions (which are actually very similar to the parameters we estimate from the full model, as they should be). Finally, we use these estimates to correct for selection bias and calculate the wage gain of going to college for college goers both in the data and as predicted by the model. We estimate a percentage gain of 72% in the data, and of 75% in the simulated data, and the difference is not significant. We take this as evidence that our factor model assumptions produce similar results than less parametric approaches like the control function.

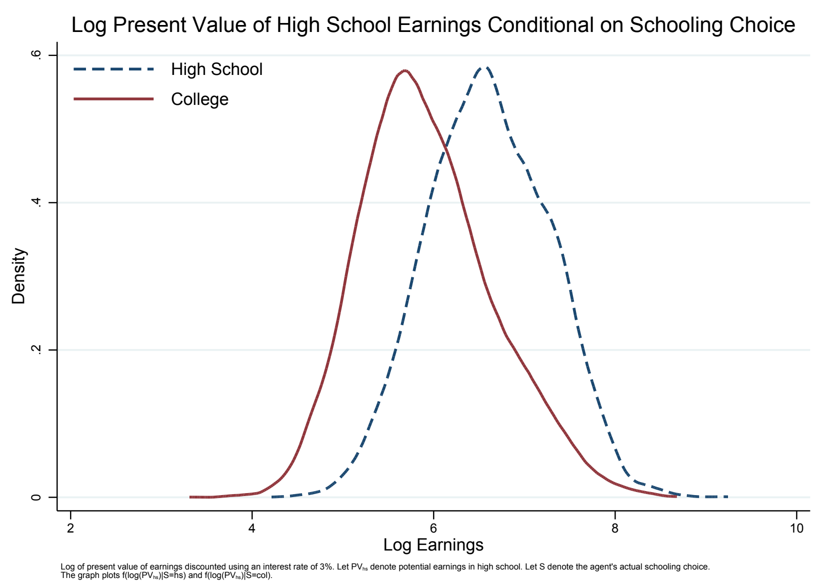


Figure 3.2: Log Present Value of High School Earnings Conditional on Schooling Choices

3.5.4 Counterfactual Analysis, Variability and Uncertainty

Figures 3.2 and 3.3 compare fitted and counterfactual distributions of log earnings for each schooling level. The figures show that people who actually stop at high school make more money as high school graduates than college graduates would have had they stopped at high school. In a similar manner, college graduates have higher college earnings than high school graduates would have had they graduated

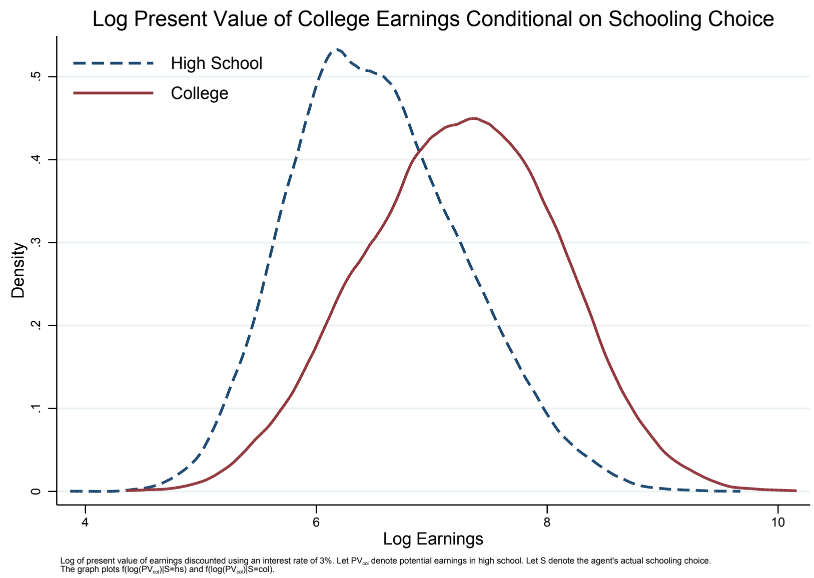


Figure 3.3: Log Present Value of College Earnings Conditional on Schooling Choices

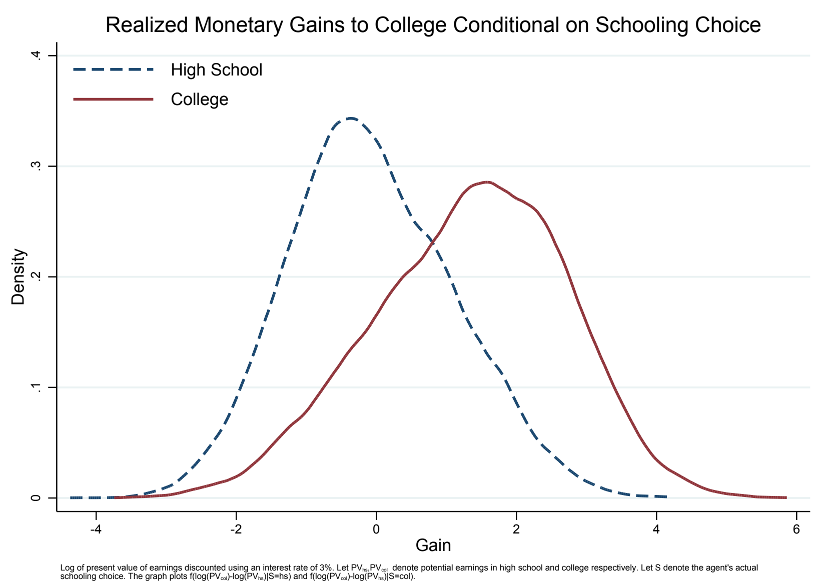


Figure 3.4: Realized Monetary Gains to College Conditional on Schooling Choice

college. So, at least in terms of earnings, people seem to be sorting in the right direction. In Figure 3.4 we plot the difference in log lifetime earnings, i.e., the ex-post realized gain of attending college, implied by the previous figures. Consistent with the sorting shown in the previous figures, the average annual return to college for high school graduates (i.e., treatment-on-the-untreated) is -1.6%, while the annual return for college graduates is 20.5%.[30] This, however, is not the whole story. Two additional features are worth noting. First, a large proportion of high school graduates would have obtained positive gains if they had gone to college; and second, a considerable fraction of college graduates get an ex-post negative gain.

If people based their decisions only on monetary gains, and observed (as opposed to expected) gains were the appropriate number to look at when explaining schooling decisions, one may be able to explain the proportion of high school graduates who would have obtained positive returns with, for example, credit constraints. One would be hard pressed to find an explanation to the significant proportion of people attending college who actually obtain a negative ex-post monetary gain. A major advantage of writing a general model of schooling decision under uncertainty is that it allows us to distinguish between expected (by the agent) and observed (by the econometrician) outcomes.

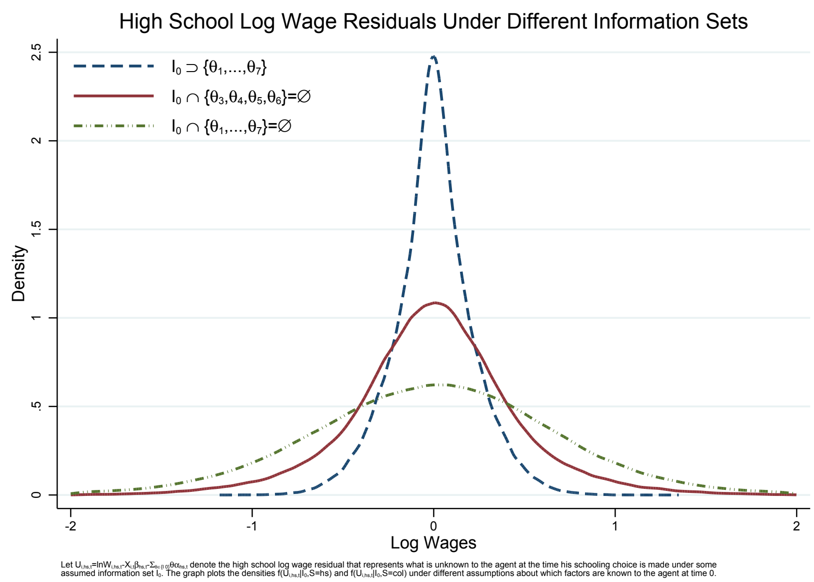


Figure 3.5: High School Log Wage Residuals Under Different Information Sets

In order to get an idea of the difference between variability and uncertainty, Figures 3.5 and 3.6 present the distributions of log wages under three different assumed information sets for the agent at the time the schooling decision is made. Even though there is a lot of dispersion, not all of it is truly uncertain to the agent. The variability in college wages is larger than that in college, and it is reduced by about the same proportion as we go from a state in which the agent knows no factor to one in which he knows $(\theta_{i,1}, \theta_{i,2}, \theta_{i,7})$ – the correctly specified information set. Figure 3.7 shows that wage gains to college are also predictable and that the

[30] All annual figures are calculated assuming it takes 4.5 years to get a college degree.

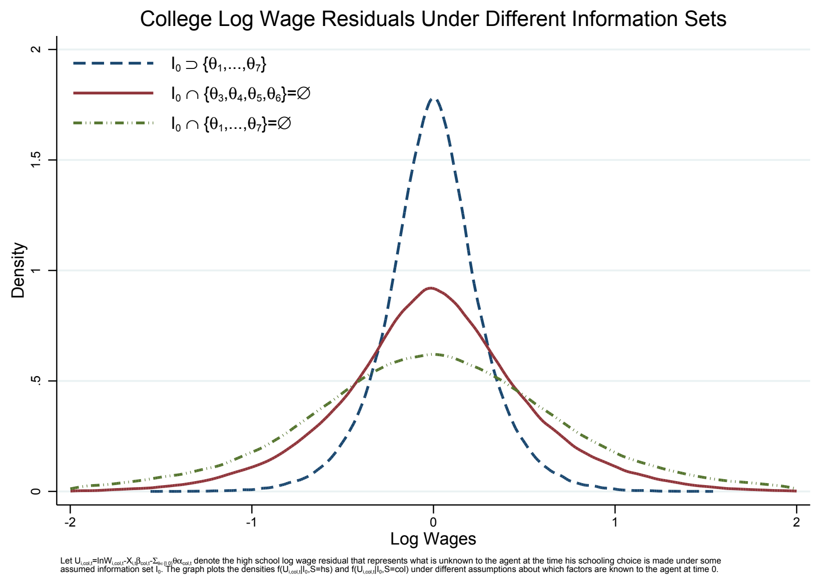


Figure 3.6: College Log Wage Residuals Under Different Information Sets

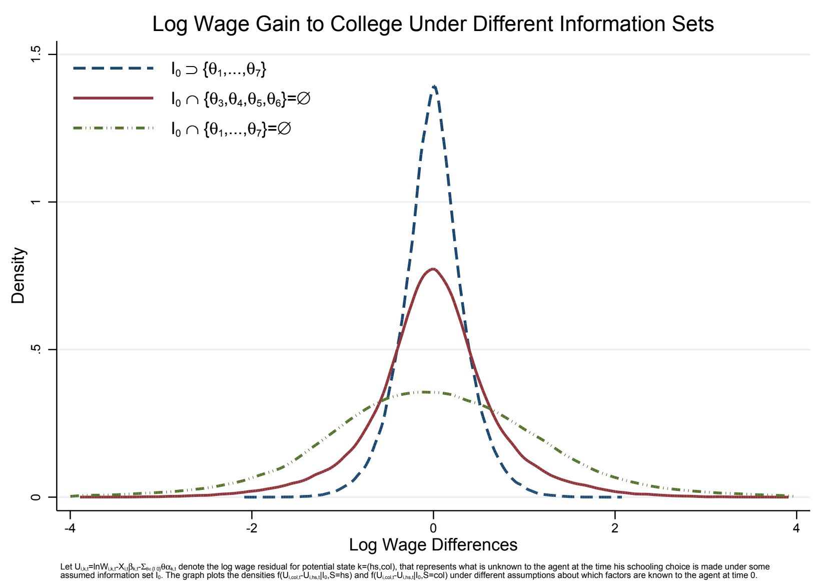


Figure 3.7: Log Wage Gain to College Under Different Information Sets

variability is reduced more than the levels are.

This pattern is more explicitly analyzed in Table 3.5. In particular, under the estimated information set at time 0 (i.e., including $\theta_{i,1}, \theta_{i,2}$ and $\theta_{i,7}$) roughly 56% –52%– of what would otherwise be considered uncertainty in high school –college– wages is predictable by the agent at the time his schooling decision is made. In the

Table 3.5: Agent's Forecast of the Variance of the Log Wages Under Different Information Sets at Schooling Choice Date

		$\textbf{Var}(\textbf{lnW}_{col})$	$\textbf{Var}(\textbf{lnW}_{hs})$	$\textbf{Var}(\textbf{lnW}_{col} - \textbf{lnW}_{hs})$
$\mathcal{I}_0^E \cap \{\theta_1, \cdots, \theta_7\} = \emptyset$ (no information)	Variance	0.5756	0.4754	1.5152
$\mathcal{I}_0 \cap \{\theta_3, \cdots, \theta_6\} = \emptyset$	Variance	0.2767	0.2092	0.4952
(estimated information set)	Fraction of Variance[2] with I_0^E explained by I_0	51.92%	56.00%	67.32%
$\mathcal{I}_0 \supset \{\theta_1, \cdots, \theta_7\}$	Variance	0.0741	0.0480	0.1221
("full" information)	Fraction of Variance[3] with I_0^E explained by I_0	87.12%	89.90%	91.94%

1. Variance of the unpredictable component of log wages as predicted at age 18.
2. The variance of the unpredictable component of high school log wages with $\mathcal{I}_0$ as given by is (1-0.5192)*0.5756=0.2767..
3. The variance of the unpredictable component of high school log wages with $\mathcal{I}_0 \supset \{\theta_1, \cdots, \theta_7\}$ is (1-0.8712)*0.5756=0.0741.

Table 3.6: Proportion of People who, when Guaranteed their Expected Wages (Keeping Credit Constraints in Place), Regret their Choice

	Choice Under Guaranteed Wage	
Choice Under Uncertainty	High School	College
	Choose HS: 51.69%	Choose Col: 48.38%
High School Choose HS: 51.69%	92.78%	7.22%
College Choose Col: 48.31%	7.58%	92.42%

1. For example, of the 51.69% of people who originally chose high school, 92.78% still do, while 7.22% now choose college. This leads to a total of 51.69% choosing high school under guaranteed wages.

same manner, the variance of the wage gains to college when the information set contains $\theta_{i,1}, \theta_{i,2}$ and $\theta_{i,7}$ is only 33% of the one we would obtain if we assumed the unobservables for the agent and the analyst coincide. That is, of the total observed variability in high school log wages, only 44% (i.e., 0.2092/0.4754) constitutes true uncertainty for the individual, i.e., variability not explained by his information. Similarly for college where 48% is left unexplained, 33% for wage gains.

Establishing this difference is clearly important for interpretation of the results. In order to study the importance of uncertainty for schooling decisions we perform the following experiment. In Tables 3.6 and 3.7 we allow people to either insure against or eliminate uncertainty in wages, *keeping the credit constraints in effect*. In order to do so, we either assume that agents can get insurance that guarantees them their expected wage given the information set at period 0 (Table 3.6),[31] or that

[31] The assumption being that the insurer gets the difference between the expected wage and the

Table 3.7: Proportion of People who, after Observing their Realized Wages (Keeping Credit Constraints in Place), Regret their Choice

	Choice Under Certainty	
Choice Under Uncertainty	High School	College
	Choose HS: 46.03%	Choose Col: 53.97%
High School		
Choose HS: 51.69%	78.84%	21.16%
College		
Choose Col: 48.31%	10.92%	89.08%

1. For example, of the 51.69% of people who originally chose high school, 78.84% still do, while 21.16% now choose college. This leads to a total of 46.03% choosing high school under certainty.

they get perfect certainty about what their wages will be in the future (Table 3.7). We keep the constraints at the original level in order to get a picture of the importance of uncertainty alone, even though the interpretation of the model implies that credit constraints arise as a consequence of uncertainty. Section 3.5.7 relaxes both.

Table 3.6 shows the importance of looking at micro evidence when accounting for uncertainty in schooling decisions. If we look only at aggregate numbers when we allow agents to insure against wage uncertainty the simulation shows essentially no effect of allowing agents to insure (keeping credit constraints in place) on college graduation. However, when we look at the micro evidence it becomes evident that the effect is much larger. Roughly 7.2% of the individuals who choose to stop at high school under uncertainty would choose to graduate college and roughly 7.6% of college graduates would regret their choice under uncertainty and would have stopped at high school.

A similar, but more significant, pattern shows up in Table 3.7 when we allow individuals to make decisions under complete certainty about wages. In the aggregate, college attendance increases by 5.7%-points, from 48.3% to 54.0%. At the individual level, we can see there is a lot of ex-post regret. Roughly 21% of the individuals who choose to stop at high school under uncertainty would choose to graduate college instead, and roughly 11% of college graduates would regret their choice under uncertainty and would have stopped at high school under complete certainty. Notice that, for both tables, the micro findings match the patterns of Figure 3.7: a proportion of high school graduates would have earned positive ex-post gains had they gone to college, and a fraction of college graduates would get negative ex-post gains and so may regret their decision ex-post. Since there is little action for the insurance case compared to the case of certainty, from now on we focus on the second one.

realized one.

Table 3.8: Average Annual Ex-post Gains and Equivalent Variations with and without Uncertainty1 (Keeping Credit Constraints in Place)

	Choice Under Uncertainty		**Choice Under Certainty**	
	High School	**College**	**High School**	**College**
Ex-post Gain1	-1.57%	20.56%	-7.04%	21.01%
Equivalent Variation2	-10.27%	30.61%	-10.64%	33.26%

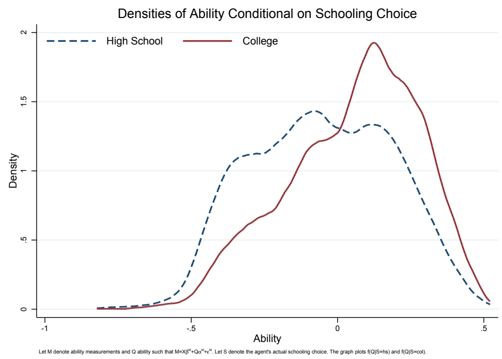
1. Certainty is defined as allowing agents to observe and choose based on their realized ex-post wages.
2. Let PV_{col} be the present value of earnings in college and PV_{hs} for high school. The lifetime ex-post gain to college is defined as $G = logPV_{col} - logPV_{hs}$. The annual ex-post gain is given by $((1+G)(1/4.5)) - 1$, under the assumption that it takes 4.5 years to obtain a college degree.
3. The lifetime equivalent variation, EV, is defined by the proportion of consumption (in each period) in the high school state that an agent is willing to give up (or needs to be compensated by) in order to be indifferent between choosing high school or college. The annual equivalent variation is given by $(1+EV)(1/4.5) - 1$, under the assumption that it takes 4.5 years to obtain a college degree.

As one would expect, there is increased sorting in terms of ex-post gains. As shown in Table 3.8, whereas the average ex-post annual gain to college for a college graduate who makes his decision using his expectations about future outcomes (i.e., integrating out the unknown $\{\theta_{i,k}\}_{k=3}^{6}, \{\varepsilon_{i,hs,t}, \varepsilon_{i,col,t}\}_{t=1}^{8}$) is 20.56%, it would increase to 21.01% if he were allowed to choose based on actual realized earnings. The same experiment shows that the average ex-post gains for high school graduates would decrease from -1.57% to -7.04%. Hence, the "best" high school graduates stop at high school, and similarly for college.

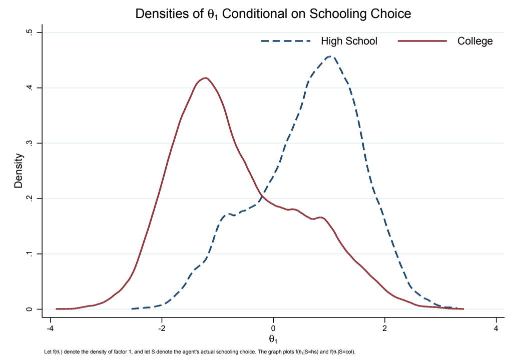
3.5.5 Sorting on Ability and Factors

In Figures 3.8(a-d) graph the densities of ability and factors 1, 2 and 7 conditioning on schooling choice. Since the remaining factors are estimated to not be known by the agent at the time the schooling decision is made, there is no selection based on them so the distribution does not change by schooling. There is strong evidence of selection in terms of ability, θ_1, and θ_7. The distribution of ability for college graduates is to the right of that for high school graduates. Individuals strongly sort in terms of ability, even after controlling for family background and individual characteristics at test date. People with higher ability graduate college more than people with lower ability. The average ability for high school graduates is 0.63 of a standard deviation lower than that of college graduates. Numerically, the difference is 0.14. A reverse pattern holds true for θ_1, θ_2 and θ_7.

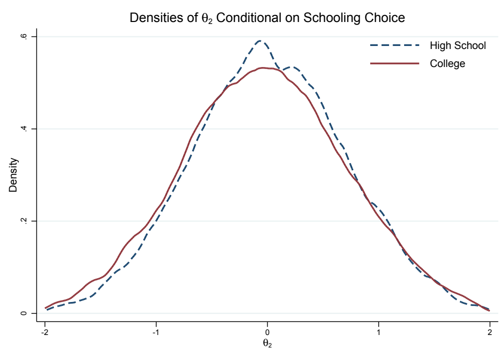
Since the factors are not tied to an external set of measurements (like ability is), it is hard to provide an interpretation for them. In order to aid in interpretation, in Table 3.9 we present the correlation of ability, $\theta_1, \theta_2, \theta_7$, and wages, hours, earnings and schooling choices. Ability is positively correlated with college wages, hours



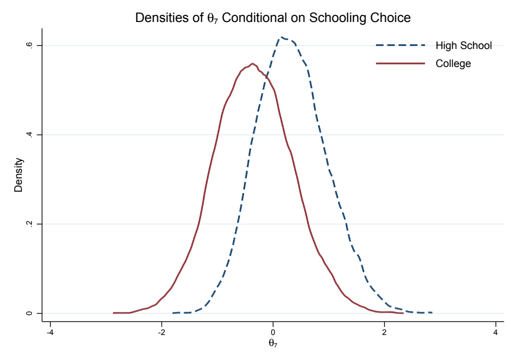
(a) Densities of Ability Conditional on Schooling Choice



(b) Densities of θ_1 Conditional on Schooling Choice



(c) Densities of θ_2 Conditional on Schooling Choice



(d) Densities of θ_7 Conditional on Schooling Choice

Figure 3.8: Density of Ability and Factors Conditional on Schooling Choice

and earnings, as well as with college attendance, but uncorrelated with their high school counterparts. θ_1, on the other hand, is positively correlated with high school wages, hours and earnings but negatively correlated with their college counterparts as well as with college attendance. θ_2 is also positively correlated with high school outcomes and negatively with college outcomes (to a lesser degree for both), but it is very weakly correlated with schooling. θ_7, on the other hand, is very weakly correlated with outcomes but highly correlated with schooling. In summary, θ_2 determines outcomes but not schooling, θ_7 determines schooling but not outcomes, while θ_1 and ability determine both.

As a consequence, only ability and θ_1 are important for explaining the difference in Figure 3.7 between high school and college graduates. That is, the difference is in big part due the different composition in terms of ability and θ_1 of the selected population via their effect on preference for schooling, and on their effects on the value of consumption for each schooling level via their effects on outcomes.

Table 3.9: Correlation Matrix

	Ability	θ_1	θ_2	θ_7	**Log Wage**		**Log Hours**		**Log Present Value Earnings**	
					High School	**College**	**High School**	**College**	**High School**	**College**
θ_1	-0.002									
θ_2	0.007	-0.010								
θ_7	-0.008	-0.003	-0.004							
HS Log Wage	-0.001	0.427	0.318	0.038						
College Log Wage	0.072	-0.404	-0.121	0.049	-0.148					
HS Log Hours	0.000	0.346	0.268	0.036	0.672	-0.116				
College Log Hours	0.055	-0.253	-0.087	0.040	-0.040	0.566	0.056			
HS Log PV Earnings	-0.002	0.770	0.436	0.034	0.629	-0.335	0.503	-0.224		
College Log PV Earnings	0.145	-0.733	-0.118	0.055	-0.334	0.556	-0.305	0.305	-0.562	
Schooling	0.228	-0.525	-0.035	-0.457	-0.239	0.230	-0.206	0.134	-0.435	0.411

3.5.6 The Importance of Preferences

From the evidence presented so far, one thing should be clear: high school graduates and college graduates are not the same. There is a great deal of heterogeneity among people. Finding high returns to college for people who actually attend college (what is commonly called "treatment on the treated" in the evaluation literature) is not necessarily informative about how much people who choose not to attend college would make if they were to attend college.

Furthermore, not all of this variability is uncertain to the agent at the time he is making his decisions. He can actually forecast a considerable proportion of it. Even though this goes a long way to explain why individuals go to college, we are still facing the question of why some people would not take advantage of the average positive gain they would get if they attended college.

So far the fact that individual decisions are based on utility maximization and not on income comparisons has been ignored. The estimated risk aversion coefficient in the model is 0.5. While this number is on the low end of the numbers reported by Browning, Hansen, and Heckman (1999), agents are in fact risk averse and do not care only about monetary returns.

A number that better summarizes the gain that individuals obtain from their schooling decisions, the "equivalent variation", is presented in the second line of Table 3.8. This number is calculated by solving for the fraction by which high school consumption needs to be changed every period for an individual to be indifferent between choosing high school or college. The equivalent variation is the consumption value that an individual places on his schooling decision accounting for preferences and, depending on the case, uncertainty. As shown in the table, on average, high school graduates would need their consumption to be reduced by 10.3% every period for them to be indifferent. That is, even though high school graduates on average face a very small monetary loss of going to college, once we account for the effect of preferences and uncertainty this is loss is much larger. In the same manner, the seemingly large 20.5% ex-post gain college graduates obtain on average is increased to 30.6% once preferences for consumption and leisure are accounted for.

In Figure 3.9 we plot the difference in the value function of attending college vs not attending, gross of the psychic cost, as perceived by the agent at the time the schooling decision is made. It is immediately apparent that, even though people who choose to go to college have a higher gross utility return to college than high school graduates, this is not enough to account for differences in college attendance. That is, we still find people who choose college with negative gross differences in the value of consumption and labor, and people who choose high school with positive differences.

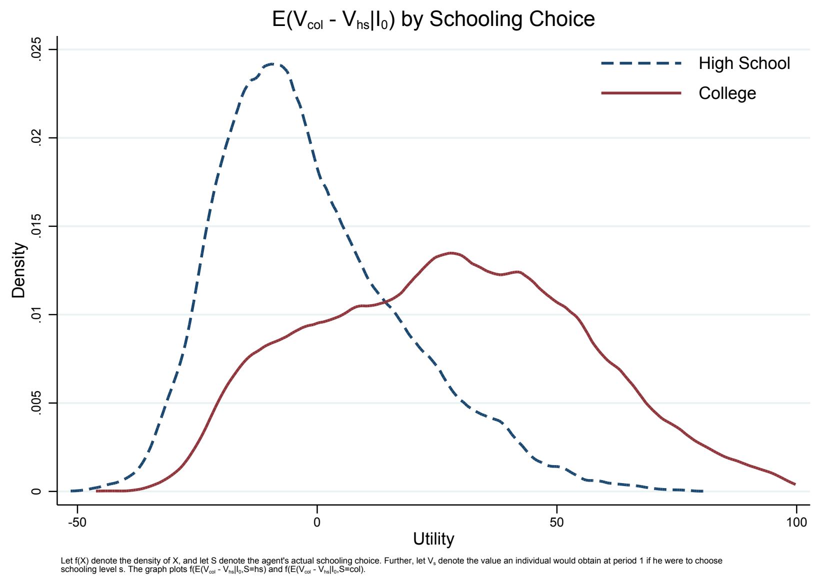


Figure 3.9: $E(\mathcal{V}_{col} - \mathcal{V}_{hs}|\mathcal{I}_0)$ by Schooling Choice

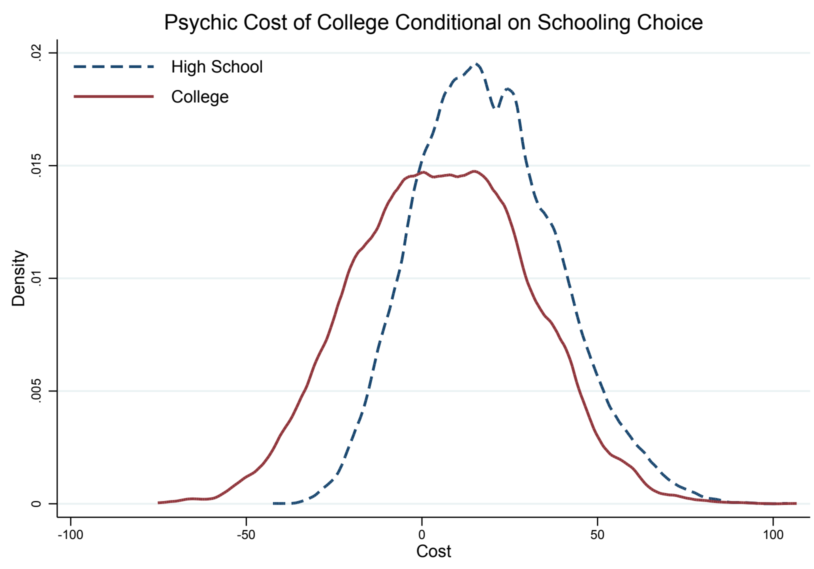


Figure 3.10: Psychic Cost of College Conditional on Schooling Choice

As shown in Figure 3.10, the remaining part is captured by the difference in preferences for schooling (the "psychic" cost function). People with low psychic costs (i.e., people who have a preference for school) are more likely to finish college. This preference is mostly driven by ability (since people with higher ability tend to have lower psychic costs) and by $\theta_{i,1}$, and to a lesser extent by $\theta_{i,7}$.

Table 3.10: Percentage of People Who Choose College Under Different Scenarios

Original Economy	48.31%
Zero Tuition Economy	48.63%
Guaranteed Expected Wages with Credit Constraints	48.38%
Certainty with Credit Constraints	53.97%
Certainty without Credit Constraints	57.87%

3.5.7 Credit constraints, risk aversion and uncertainty

The final step in our analysis of schooling decisions is to account for credit constraints. For this purpose, we perform two different simulations. First, we follow the literature on the effect of borrowing constraints on schooling and look at the inability of people to pay for tuition. Our first simulation consists on setting tuition to zero for everyone. The results are shown in the second line of Table 3.10. Making college free for everyone effectively relaxes the credit constraint since it allows people to increase their consumption by the amount of money that they would otherwise dedicate to tuition. It also captures the effect of reducing the price of schooling. As a consequence, the result of this exercise is an upper bound on the effect on relaxing the constraint via the tuition subsidy. There is a increase of 0.3%-points in college attendance. The result is consistent, although even smaller, with the findings of Keane and Wolpin (2001), Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2006) where a similar experiment is performed and they all find a small effect. In the second simulation, borrowing constraints are eliminated by allowing people to select using their realized earnings. The effects of eliminating the credit constraints and uncertainty with them are significantly different from the ones obtained from just eliminating tuition. For completeness, in the third and fourth lines of table 3.10 we repeat the exercise of reducing uncertainty while keeping the credit constraints in place. We can see that college attendance increases very little when agents are allowed to insure against wage risk, while the increase is larger (5.7%-points) when uncertainty is eliminated completely. In the fifth line we eliminate both uncertainty and the borrowing constraints. In this case, we are back to a standard permanent income model. Notice that people are able to smooth consumption much more effectively under this setting. For this case, the increase in college attendance is much more substantial (almost 10%-points), to roughly 57.8% when consumption smoothing is allowed. Combining this with the fourth line in the table, we conclude that the "pure" effect of eliminating the credit constraint is to increase college attendance by 3.9%-points. As compared to the case in which only uncertainty is eliminated, by letting people smooth consumption by eliminating the borrowing constraints, their consumption at young ages (while in college) need not decrease as much.

Although these numbers clearly point to both uncertainty and credit constraints playing an important role, they should be interpreted with some caution. At least two of the assumptions we make contribute to these results. First, we are assuming that all risks are idiosyncratic. If some risks were aggregate (or in general not all risks were insurable) clearly, people would not be able to perfectly smooth consumption and the result would be less dramatic. The tuition subsidy example shows that this is the case by relaxing credit constraints but not completely eliminating them as would be the case of aggregate shock. Finally, as opposed to the case in which credit constraints are slightly relaxed via tuition, we now completely change the way the economy operates. General equilibrium effects would certainly dampen this response.[32] Even with the caveats just mentioned, the effect of credit constraints seems large enough to play an important role. If one were to cut the effect in half so that now college graduation only increases from 48.3% to 53.1% it is still much larger than the effect obtained by setting tuition to zero.

3.6 Conclusion

This paper analyzes the role played by ability, uncertainty, preferences and credit constraints in explaining schooling choices. The conclusion of the paper is that there is no clear candidate for "the" explanation as to why some people got to college and some don't. This may not be surprising result given the nature of the decision, but it is by no means obvious a priori. As the results in this paper show, the college attendance decision is composed of many parts and all of them help explain the patterns in the data. Ability, preferences and uncertainty all play important roles. Once borrowing constraints are clearly defined and people are allowed to smooth consumption, as opposed to simply relaxing the constraint, credit constraints play a more important role than previously found.

To the extent that simpler models are easier to interpret and require weaker assumptions they are preferable. However, they cannot always answer the questions one maybe interested in. In very complex models, however, the relationship between results, assumptions and data is hard to visualize. In an effort to show that the results in this paper are not simply a consequence of functional forms or distributional assumptions, we prove that the model we use is semiparametrically identified.

We build on the work Carneiro, Hansen, and Heckman (2003) and Cunha, Heckman, and Navarro (2005) that identifies the uncertainty facing the agent. We find that agents can predict a considerable proportion of the variance of their future earnings as well as their gains at the time schooling decisions are made. The remaining uncertainty helps explain why some people do not go to college even though they would obtain positive returns and why some individuals attend college when their observed ex-post return is negative. Individuals make their decisions

[32]See Heckman, Lochner, and Taber (1998a,b, 1999) for evidence on how important this general equilibrium effects can be.

before all relevant information is revealed.

Ability is an important determinant of schooling choice. Individual ability helps explain college attendance mostly through the individual preference for school. High ability individuals have lower psychic costs of attending college and it is mainly through this channel that ability affects college attendance.

We find little evidence of liquidity constraints when these are defined as individuals not being able to afford college. Moving to an economy with zero tuition increases college attendance by roughly 0.3%-points. When both borrowing constraints and uncertainty are eliminated on the other hand, the effect is much larger and now college attendance increases by almost 10%-points. Once credit constraints are defined in terms of consumption smoothing, instead of liquidity constraints at a point in time, they play a stronger role than previously found. Our simulation of a feasible uncertainty reducing policy that guarantees agents their expected earnings, however, shows very little impact.

Bibliography

Aguirregabiria, Victor and Pedro Mira. 2002. "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models." *Econometrica* 70 (4):1519–1543.

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *Quarterly Journal of Economics* 109 (3):659–684.

Ashenfelter, Orley and Alan B. Krueger. 1994. "Estimates of the Economic Returns to Schooling from a New Sample of Twins." *American Economic Review* 84 (5):1157–1173.

Attanasio, Orazio and Hamish Low. 2004. "Estimating Euler Equations." *Review of Economic Dynamics* 7:405–435.

Belzil, Christian and Jorgen Hansen. 2002. "Unobserved Ability and the Return to Schooling." *Econometrica* 70 (5):2075–2091.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2004. "Consumption inequality and partial insurance." Tech. Rep. WP04/28, Institute for Fiscal Studies.

Blundell, Richard and Ian Preston. 1998. "Consumption Inequality and Income Uncertainty." *Quarterly Journal of Economics* 113 (2):603–640.

Brown, Meta, John Karl Scholz, and Ananth Seshadri. 2012. "A New test of Borrowing Constraints for Education." *Review of Economic Studies* 79 (2):511–538.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. "Micro Data and General Equilibrium Models." In *Handbook of Macroeconomics*, vol. 1A, edited by John B. Taylor and Michael Woodford, chap. 8. Elsevier, 543–633.

Browning, Martin and Annamaria Lusardi. 1996. "Household Saving: Micro Theories and Micro Facts." *Journal of Economic Literature* 34 (4):1797–1855.

Cameron, Stephen V. and James J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy* 106 (2):262–333.

———. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109 (3):455–99.

Cameron, Stephen V. and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *Journal of Political Economy* 112 (1):132–182.

Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5):1127–1160.

Carneiro, Pedro, Karsten Hansen, and James J. Heckman. 2003. "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice." *International Economic Review* 44 (2):361–422. 2001 Lawrence R. Klein Lecture.

Carneiro, Pedro and James J. Heckman. 2002. "The Evidence on Credit Constraints in Post-Secondary Schooling." *Economic Journal* 112 (482):705–734.

Carranza, Juan Esteban. 2007. "Product innovation and adoption in market equilibrium: The case of digital cameras." Unpublished manuscript, University of Wisconsin-Madison, Department of Economics.

Cawley, John, James J. Heckman, Lance J. Lochner, and Edward J. Vytlacil. 2000. "Understanding the Role of Cognitive Ability in Accounting for the Recent Rise in the Return to Education." In *Meritocracy and Economic Inequality*, edited by K. Arrow and S. Bowles. Princeton, NJ: Princeton University Press.

Chioda, Laura. 2004. "Estimating Euler Equations with Measurement Error: A Nonparametric Approach." Unpublished Manuscript, University of California Berkeley, Department of Economics.

Cooley Fruehwirth, Jane, Salvador Navarro, and Yuya Takahashi. 2016. "How The Timing of Grade Retention Affects Outcomes: Identification and Estimation of Time-Varying Treatment Effects." *Journal of Labor Economics* 3 (4).

Cunha, Flavio, James J. Heckman, and Salvador Navarro. 2005. "Separating Uncertainty from Heterogeneity in Life Cycle Earnings, The 2004 Hicks Lecture." *Oxford Economic Papers* 57 (2):191–261.

———. 2006. "Counterfactual Analysis of Inequality and Social Mobility." In *Mobility and Inequality: Frontiers of Research in Sociology and Economics*, edited by Stephen L. Morgan, David B. Grusky, and Gary S. Fields, chap. 4. Stanford, CA: Stanford University Press, 290–348.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. "An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics." *Journal of Human Resources* 33 (2):251–299.

Flavin, Marjorie A. 1981. "The Adjustment of Consumption to Changing Expectations About Future Income." *Journal of Political Economy* 89 (5):974–1009.

Gourinchas, Pierre-Olivier and Jonathan Parker. 2002. "Consumption over the LifeCycle." *Econometrica* 70 (1):47–89.

Guvenen, Fatih. 2007. "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *The American Economic Review* 97 (3):687–712.

Hansen, Karsten T., James J. Heckman, and Kathleen J. Mullen. 2004. "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* 121 (1-2):39–98.

Hansen, Lart Peter and Kenneth J. Singleton. 1983. "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns." *Journalof Political Economy* 91 (2):249–265.

Heckman, James J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35 (1):1–97.

Heckman, James J., Lance J. Lochner, and Christopher Taber. 1998a. "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* 1 (1):1–58.

———. 1998b. "General-Equilibrium Treatment Effects: A Study of Tuition Policy." *American Economic Review* 88 (2):381–386.

———. 1999. "General-Equilibrium Cost-Benefit Analysis of Education and Tax Policies." In *Trade, Growth and Development: Essays in Honor of T.N. Srinivasan*, edited by G. Ranis and L.K. Raut, chap. 14. Elsevier Science B.V., 291–349.

Heckman, James J. and Salvador Navarro. 2007. "Dynamic Discrete Choice and Dynamic Treatment Effects." *Journal of Econometrics* 136 (2):341–396.

Heckman, James J. and Edward J. Vytlacil. 1998. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return Is Correlated with Schooling." *Journal of Human Resources* 33 (4):974–987.

Hotz, V. Joseph and Robert A. Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *Review of Economic Studies* 60 (3):497–529.

Kane, Thomas J. 1996. "College Cost, Borrowing Constraints and the Timing of College Entry." *Eastern Economic Journal* 22 (2):181–194.

Kaufmann, Katja and Luigi Pistaferri. 2009. "Disentangling Insurance and Information in Intertemporal Consumption Choices." *American Economic Review* 99 (2):387–92.

Keane, Michael P. and Kenneth I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3):473–522.

———. 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review* 42 (4):1051–1103.

Laitner, John. 1992. "Random Earnings Differences, Lifetime Liquidity Constraints, and Altruistic Intergenerational Transfers." *Journal of Economic Theory* 58 (2):135–170.

Lochner, Lance J. and Monge-Naranjo Alexander. 2011. "The Nature of Credit Constraints and Human Capital." *American Economic Review* 101 (6):2487–2529.

———. 2012. "Credit Constraints in Education." *Annual Review of Economics* 4:225–256.

MaCurdy, Thomas E., Thomas Mroz, and R. Mark Gritz. 1998. "An Evaluation of the National Longitudinal Survey on Youth." *Journal of Human Resources* 33 (2):345–436.

Matzkin, Rosa L. 2003. "Nonparametric Estimation of Nonadditive Random Functions." *Econometrica* 71 (5):1339–1375.

Navarro, Salvador. 2008. "Control Function." In *The New Palgrave Dictionary of Economics.*, edited by Steven N. Durlauf and Lawrence E. Blume. London: Palgrave Macmillan Press, second ed.

Peterman, William B. 2016. "RECONCILING MICRO AND MACRO ESTIMATES OF THE FRISCH LABOR SUPPLY ELASTICITY." *Economic Inquiry* 54 (1):100–120. URL `http://dx.doi.org/10.1111/ecin.12252`.

Pistaferri, Luigi. 2001. "Superior Information, Income Shocks, and the Permanent Income Hypothesis." *Review of Economics and Statistics* 83 (3):465–476.

Roehrig, Charles S. 1988. "Conditions for Identification in Nonparametric and Parametric Models." *Econometrica* 56 (2):433–447.

Sims, Christopher A. 1972. "Money, Income, and Causality." *American Economic Review* 62 (4):540–552.

Taber, Christopher R. 2001. "The Rising College Premium in the Eighties: Return to College or Return to Unobserved Ability?" *Review of Economic Studies* 68 (3):665–691.

Ventura, Eva. 1994. "A Note on Measurement Error and Euler Equations." *Economics Letters* 45:305 – 308.

Bibliography

Banerjee, B. 1984. “Information Flow, Expectations and Job Search: Rural-to Urban Migration Process in India.” *Journal of Development Economics* 15:239–257.

Banerjee, Biswajit. 1983. “Social Networks in the Migration Process: Empirical Evidence on Chain Migration in India.” *Journal of Development Areas* 17:185–196.

Beringhaus, S. and H.G. Seifer-Vogt. 1988. “Temporary versus Permanent Migration: a Decision Theoretical Approach.” *Journal of Population Economics* 1:195–211.

Blejer, M.I. and I. Goldberg. 1980. “Return Migration-expectations versus Reality: a Case Study of Western Immigrants to Israel.” *Research in population economics* 2.

Borjas, G. and B. Bratsberg. 1996. “Who leaves? The Outmigration of the Foreign-born.” *Review of Economics Statistics* 78:165–176.

Borjas, G.J. 1999. *The Economic Analysis of Immigration.*, vol. 3A. Amsterdam: North-Holland.

Burdett, Kiefer N.M. Mortensen D.T., K. and G.R. Neumann. 1984. “Earnings, unemployment, and the allocation of time over time.” *Review of Economic Studies* 51(4):559–78.

Chiswick, B.R. 1978. “The effect of Americanization on the Earning of Foreign-born Men.” *Journal of Political Economy* 86:897–921.

Flinn, C. and J. Heckman. 1982. *Models for the Analysis of Labour Force Dynamics*. Greenwich: JAI Press.

Frijters, P and Bas van der Klaauw. 2006. “Job Search with Nonparticiparion.” *The Economic Journal* 116:45–83.

Galor, O. and O. Stark. 1990. “Migrants’ Savings, the Probability of Return Migration, and Migrants.” *International Economic Review* 31:463–467.

Hare, Denise and Shukai Zhao. 2000. “Lobor Migration as a Rural Development Strategy: a View from the Migration Origin.” *Working paper* .

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

Heckman, James and B. Singer. 1984. "A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52:271–320.

Huang, P. and F.N. Pieke. 2003. "China Migration Country Study."

Jasso, G. and M.R. Rosenzweig. 1982. "Estimating the Emigration Rates of Legal Immigrants Using Administrative and Survey Data: the 1971 Cohort of Immigrants to the US." *Demography* 19:279–290.

Kiefer, N.M. and G. Neuman. 1979. *Studies in the Economics of Search*, chap. Estimation of Wage Offer Distributions and Reservation Wages. Amsterdam: North-Holland.

Lin, F. Cai, Justin. and Z. Li. 1996. *The China Miracle: Development Strategy and Economic Reform*. Hong Kong: Chinese University Press.

Machin, S. and A. Manning. 1999. *Handbook of Labor Economics*, vol. 3C, chap. The Cause and Consequences of Longterm Unemployment in Europe. Amsterdam: North-Holland, 3085–139.

McCall, B. 1996. "Unemployment Insurance Rules, Joblessness, and Part-time Work." *Econometrica* 64:647–682.

Meng, Xin. 2000. *Labout Market Reform in China*. Cambridge University Press.

Mortensen, D.T. 1986. "Job Search and Labor Market Analysis." *Handbook of Labor Economics* 2:849–919.

Mortensen, D.T. and G.R. Neuman. 1984. *Choice or Chance? A Structural Interpretation of Individual Labor Market Dynamics*. Berlin: Springer-Verlag, 98–131.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labour Market." *Quarterly Journal of Economics* 118(2):549–599.

Murphy, Rachel. 1999. "Return Migration Entrepreneurs and Economic Diversification in Two Counties in South Jiangxu, China." *Journal of International Development* 11:662–672.

Todaro, M. P. 1969. "A Model of Labor Migration and Urban Unemployment in Less Development Countries." *American Economic Review* 59:138–48.

Uhlendorff, A. and K.F. Zimmermann. 2006. "Unemployment dynamics among Migrants and Natives." *Working paper* .

Van den Berg, G. 2001. “Duaration Models: Specification, identication, and multiple duration.” *Handbook of Econometrics* 5:3381–460.

Van den Berg, G.J. 1990. “Search Behavior, Transitions to Nonparticiparion and the Duration of Unemployment.” *Economic Journal* 100:842–65.

Weiner, S.E. 1984. *A Survival Analysis of Adult Male Black-White Labor Market Flows*. Berlin: Springer-Verlag, 132–57.

Wolpin, K.I. 1987. “Estimating a Structural Search Model: the Transition From School to Work.” *Econometrica* 55(4):801–17.

Wu, Harry Xiaoying. 1994. “Rural to Urban Migration in the People’s Republic of China.” *China Quarterly* 139:669–98.

Yao, Y. 2001. “Egalitarian Land Distribution and Labor Migration in Rural China.” *working paper* .

Zhang, Junfu and Zhong Zhao. 2011. “Social-Family Network and Self- Employment: Evidence from Temporary Rural-Urban Migrants in China.” *working paper* .

Zhao, Yaohui. 1997. “Labor Migration and Returns to Rural Education in China.” *American Journal of Agricultural Economics* 79:1278–87.

———. 2000. “Rural-to-Urban Labor Migration in China: the Past and the Present.” .

———. 2002. “Causes and Consequences of Return Migration: Recent Evidence from China.” *Journal of Comparative Economics* 30:376–394.

———. 2003. “The Role of Migrant Networks in Labour Migration: the Case of China.” *Contemporary Economic Policy* 21(4):500–511.

Zhao, Zhong. 2005. “Migration, Labor Market Flexibility, and Wage Determination in China: A Review.” *Developing Economics* XLIII-2:285–312.

Appendix A

Appendix for Chapter 1: Theorem

A.1 The proof of Theorem 1.3.1

Bellman Equation is:

$$V_u = c + \frac{\lambda}{1+r} E_{w,\eta} \max\{V_e, V_u, V_r\} + \frac{1-\lambda}{1+r} E_{\eta} \max\{V_u, V_r\} \tag{A.1}$$

The term of $E_{w,\eta} \max\{V_e, V_u, V_r\}$ can be extended into four terms:

$$\begin{aligned} E_{w,\eta} \max\{V_e, V_u, V_r\} \quad = \quad & E_{w,\eta}(V_e(1_{\{}V_e > V_u > V_r\} + 1_{\{}V_e > V_r > V_u\}) + V_u(1_{\{}V_u > V_e > V_r\} \\ & +1_{\{}V_u > V_r > V_e\}) + V_r(1_{\{}V_r > V_e > V_u\} + 1_{\{}V_r > V_u > V_e\})) \end{aligned}$$

if $V_u > 0$

$$\begin{aligned} V_1 &= E_{w,\eta}((V_e - V_r)1_{\{}V_e - V_r > V_u - V_r > 0\}) \\ &= E_{w,\eta}((\frac{w}{r} - \frac{\eta}{r})1_{\frac{w}{r} - \frac{\eta}{r} > V_u - \frac{\eta}{r} > 0}) \\ &= \int_{-\infty}^{rV_u} \int_{rV_u}^{\infty} (\frac{w-\eta}{r}) f(w) dw g(\eta) d\eta \\ &= \int_{-\infty}^{rV_u} [V_u(1 - F(rV_u)) + \int_{rV_u}^{\infty} \frac{1}{r}(1 - F(w))dw - \frac{\eta}{r}(1 - F(rV_u))] g(\eta) d\eta \\ &= [V_u(1 - F(rV_u)) + \int_{rV_u}^{\infty} \frac{1}{r}(1 - F(w))dw] G(rV_u) - \frac{1 - F(rV_u)}{r} \int_{-\infty}^{rV_u} \eta g(\eta) d\eta \end{aligned}$$

$$\begin{aligned} V_2 &= E_{w,\eta}((V_e - V_r)1_{\{}V_e - V_r > 0 > V_u - V_r\}) \\ &= E_{w,\eta}((\frac{w}{r} - \frac{\eta}{r})1_{\frac{w}{r} - \frac{\eta}{r} > 0 > V_u - \frac{\eta}{r}}) \\ &= \int_{rV_u}^{\infty} [\frac{1}{r} \int_{\eta}^{\infty} (1 - F(w))dw] g(\eta) d\eta \\ &= \int_{rV_u}^{\infty} [(V_u - \frac{w}{r}) G(rV_u) + \frac{1}{r} \int_{rV_u}^{w} G(\eta) d\eta] f(w) dw \end{aligned}$$

$$\begin{aligned} V_3 &= E_{w,\eta}((V_u - V_r)1_{\{}V_u - V_r > 0 > V_e - V_r\}) \\ &= \int_0^{rV_u}\int_w^{rV_u}(V_u - \frac{\eta}{r})g(\eta)d\eta f(w)dw \\ &= \int_0^{rV_u}(V_u(G(rV_u) - G(w)) - \frac{1}{r}\int_w^{rV_u}\eta g(\eta)d\eta)f(w)dw \\ &= \int_0^{rV_u}\int_0^{\eta}(V_u - \frac{\eta}{r})f(w)dwg(\eta)d\eta \end{aligned}$$

$$\begin{aligned} V_4 &= E_{w,\eta}((V_u - V_r)1_{\{}V_u - V_r > V_e - V_r\} > 0) \\ &= E_{w,\eta}((V_u - \frac{\eta}{r})1_{V_u - \frac{\eta}{r} > \frac{w-\eta}{r} > 0}) \\ &= \int_0^{rV_u}\int_{-\infty}^{w}(V_u - \frac{\eta}{r})g(\eta)d\eta f(w)dw \\ &= \int_0^{rV_u}[V_uG(w) - \int_{-\infty}^{w}\frac{\eta}{r}g(\eta)d\eta]f(w)dw \\ &= \int_{-\infty}^{rV_u}\int_{\eta}^{rV_u}(V_u - \frac{\eta}{r})f(w)dwg(\eta)d\eta \\ &= V_uF(rV_u)G(rV_u) + \int_{-\infty}^{rV_u}[\frac{\eta}{r}(F(\eta) - F(rV_u)) - V_uF(\eta)]g(\eta)d\eta \end{aligned}$$

The third term of Bellman Equation is:

$$\begin{aligned} E_\eta max\{V_u, V_n\} &= \int_{rV_u}^{\infty}(\frac{\eta}{r}g(\eta))d\eta + \int_{-\infty}^{V_u}V_ug(\eta)d\eta \\ &= \int_{rV_u}^{\infty}(\frac{\eta}{r}g(\eta))d\eta + V_uG(rV_u) \end{aligned}$$

The main idea of proof is to show the function $\phi(V_u)$ is monotonic. If the function $\phi(V_u)$ is strictly monotonic, then the function will have at most one solution. Then we will talk about the condition to make sure the unique solution exists.

$$\begin{aligned} \frac{dV_1}{dV_u} &= rg(rV_u)[V_u(1 - F(rV_u))] - rf(rV_u)V_uG(rV_u) + f(rV_u)\int_{-\infty}^{rV_u}\eta g(\eta)d\eta \\ &\quad +(\int_{rV_u}^{\infty}(1 - F(w))dw)g(rV_u) \end{aligned}$$

$$\frac{dV_2}{dV_1} = -(\int_{rV_u}^{\infty}(1 - F(w))dw)g(rV_u)$$

$$\frac{dV_3}{dV_u} = \int_0^{rV_u}F(\eta)g(\eta)d\eta$$

$$\frac{dV_4}{dV_u} = \int_0^{rV_u} G(w)f(w)dw + V_u G(rV_u)f(rV_u) - (\int_{-\infty}^{rV_u} \eta g(\eta)d\eta)f(rV_u)$$

$$\frac{V_1 + V_2 + V_3 + V_4}{dV_u} = rg(rV_u)(V_u(1 - F(rV_u))) + G(rV_u)F(rV_u)$$

$$\begin{aligned}
\frac{d\phi(V_u)}{dV_u} &= \frac{\lambda\{rg(rV_u)(V_u(1 - F(rV_u))) + G(rV_u)F(rV_u)\}}{1 + r} + \frac{1 - \lambda}{1 + r}G(rV_u) - 1 \\
&= \frac{\lambda(F(rV_u) - 1)(G(rV_u) - rV_u g(rV_u)) + G(rV_u) - (1 + r)}{1 + r}
\end{aligned}$$

We can prove the function $G(x) - xg(x) > 0$ if $G(x)$ is CDF and $g(x)$ is density of standard normal distribution, so $\frac{d\phi(V_u)}{dV_u} < 0$. when $V_u > 0$, the function:

$$\phi(V_u) = c + \frac{\lambda}{1 + r}E_{w,\eta} \max\{V_e, V_u, V_r\} + \frac{1 - \lambda}{1 + r}E_\eta \max\{V_u, V_n\} - V_u$$

is monotonic decreasing.

if $V_u \leq 0$

$$\begin{aligned}
V_1 &= E_{w,\eta}((V_e - V_r)1_{\{}V_e - V_r > V_u - V_r > 0\}) \\
&= E_{w,\eta}((\frac{w}{r} - \frac{\eta}{r})1_{\frac{w}{r} - \frac{\eta}{r} > V_u - \frac{\eta}{r} > 0}) \\
&= \int_{-\infty}^{rV_u} \int_0^\infty (\frac{w - \eta}{r})f(w)dwg(\eta)d\eta \\
&= \int_{-\infty}^{rV_u} \{\frac{1}{r} \int_0^\infty (1 - F(w))dw - \frac{\eta}{r}(1 - F(0))\}g(\eta)d\eta
\end{aligned}$$

$$\begin{aligned}
V_2 &= E_{w,\eta}((V_e - V_r)1_{\{}V_e - V_r > 0 > V_u - V_r\}) \\
&= E_{w,\eta}((\frac{w}{r} - \frac{\eta}{r})1_{\frac{w}{r} - \frac{\eta}{r} > 0 > V_u - \frac{\eta}{r}}) \\
&= \int_0^\infty \int_{rV_u}^w (\frac{w - \eta}{r})g(\eta)d\eta f(w)dw \\
&= \int_0^\infty \frac{w}{r}G(w)f(w)dw - \int_0^\infty \frac{w}{r}G(rV_u)f(w)dw - \int_0^\infty (\int_{rV_u}^w \frac{\eta}{r}g(\eta)d\eta)f(w)dw
\end{aligned}$$

$$\frac{dV_1}{dV_u} = (\int_0^\infty (1 - F(w)))g(rV_u) - rV_u g(rV_u)(1 - F(0))$$

$$\frac{dV_2}{dV_u} = -rg(rV_u) \int_0^\infty \frac{w}{r}f(w)dw + rV_u g(rV_u) \int_0^\infty f(w)dw$$

$$\frac{d\phi(V_u)}{dV_u} = \frac{1 - \lambda}{1 + r}G(rV_u) - 1 < 0$$

Above proof shows $\phi(V_u)$ is strictly monotonic function. Denote $h(V_u)$ as:

$$h(V_u) = \frac{\lambda}{1+r} E_{w,\eta} \max\{V_e, V_u, V_r\} + \frac{1-\lambda}{1+r} E_\eta \max\{V_u, V_r\} - V_u$$

$\because$ E max is an increasing operator so $h(V_u) \geq 0$;
$\because$ $h(V_u) = \phi(V_u) - c$ and $\phi(V_u)$ is strictly decreasing function of V_u,
$\therefore$ $h(V_u)$ is strictly decreasing function of V_u.

The value space of $h(V_u) = [0, \infty)$ and $\lim_{V_u \to -\infty} h(V_u) = \infty$;
$\lim_{V_u \to \infty} h(V_u) = 0$.
$\therefore$ For $\forall$ c, s.t. $c < 0$ $\exists!$ V_u s.t.

$$h(V_u) = -c$$

$\because$ $h(V_u) = \phi(V_u) - c$,
$\therefore$ $\phi(V_u) = 0$ and $\exists!$ V_u s.t $\phi(V_u) = 0$ ■

Appendix B

Appendix for Chapter 1: Likelihood

B.1 Likelihood Function

$$
\begin{aligned}
\Pr 1 &= \Pr(V_e > \max\{V_u, V_r\}) \\
&= \Pr(V_e > V_u > V_r) + \Pr(V_e > V_r > V_u) \\
&= 1_{V_u \geq 0}(\int_{-\infty}^{rV_u} \int_{rV_u}^{\infty} f(w)dwg(\eta)d\eta + \int_{rV_u}^{\infty} \int_{rV_u}^{w} g(\eta)d\eta f(w)dw) \\
&\quad +1_{V_u<0}(\int_{-\infty}^{rV_u} \int_{0}^{\infty} f(w)dwg(\eta)d\eta + \int_{0}^{\infty} \int_{rV_u}^{w} g(\eta)d\eta f(w)dw)
\end{aligned}
$$

$$
\begin{aligned}
\Pr 2 &= \Pr(V_n > \max\{V_e, V_r\}) \\
&= \Pr(V_r > V_e > V_u) + \Pr(V_r > V_u > V_e) \\
&= 1_{V_u \geq 0}(\int_{rV_u}^{\infty} \int_{rV_u}^{\eta} f(w)dwg(\eta)d\eta + \int_{rV_u}^{\infty} \int_{0}^{rV_u} f(w)dwg(\eta)d\eta) \\
&\quad +1_{V_u<0}(\int_{0}^{\infty} \int_{0}^{\eta} f(w)dwg(\eta)d\eta)
\end{aligned}
$$

$$\Pr 3 = \Pr(V_r > V_u) = \Pr(\eta > rV_u) = \int_{rV_u}^{\infty} g(\eta)d\eta$$

Likelihood for i, Employed:

$$\lambda \Pr 1 \cdot \exp(-(\lambda \Pr 1 + \lambda \Pr 2 + (1-\lambda) \Pr 3)t)$$

Likelihood for i, Returned:

$$(\lambda \Pr 2 + (1-\lambda) \Pr 3) \exp(-(\lambda \Pr 1 + \lambda \Pr 2 + (1-\lambda) \Pr 3)t)$$

Likelihood for i, Unemployed:

$$\exp(-(\lambda \Pr 1 + \lambda \Pr 2 + (1-\lambda) \Pr 3)t)$$

Appendix C

Estimation and Simulation Results

C.1 Estimates for Marriage and Fertility Process

Table C.1: The Estimates of Marriage and Fertility Transition Process

	Marriage		Fertility
Education (β_1^{ma})	0.0027	Education (β_5^f)	-0.0062
	(0.0007)		(0.0029)
Constant (β_0^{ma})	3.1671	Age (β_1^f)	0.2709
	(0.0062)		(0.0098)
lnγ	-2.4694	Age^2 (β_2^f)	-0.0050
	(0.0096)		(0.0002)
		Num of Children$_{t-1}$ (β_3^f)	-1.1181
			(0.0200)
		Num of Children$_{t-1}^2$ (β_4^f)	0.197
			(0.0068)
		Married (β_6^f)	1.9646
			(0.0248)
		Constant (β_0^f)	-6.0428
			(0.1435)

1. The variable of education is education year.
2. Married is the indicator variable for marital status.
3. The first two columns give the estimates of the survival analysis of marital transition.
4. Column 4 give the estimates of fertility transition which is formed by a dynamic probit model.
5. Numbers in parentheses are standard errors.

Table C.2: Estimation Results for the Model without Network Investment Decision

Earning Equation(Urban)		Social Network Probit Equation	
edu year	0.0143	marriage	0.0883
	(0.0001)		(0.0001)
expu	0.0040	num of children	-0.0364
	(0.0001)		(0.0001)
expr	0.0035	sn_{t-1}	4.2777
	(0.0001)		(0.0001)
$expu^2 \times 100$	-0.0018	constant	-1.9191
	(0.0001)		(0.0001)
$expr^2 \times 100$	0.0001	Psychic value	
constant	6.4450	age	0.0369
	(0.0001)		(0.0001)
Earning Equation(Rural)		$age^2 \times 100$	-0.0018
			(0.0001)
edu year	0.0087	marriage	0.0013
	(0.0001)		(0.0001)
expu	0.0018	num of children	-0.0870
	(0.0001)		(0.0001)
expr	0.0057	constant	0.0854
	(0.0001)		(0.0001)
$expu^2 \times 100$	-0.0037	Job arrival rate	
	(0.0001)		
$expr^2 \times 100$	-0.0005	social network	0.7409
	(0.0001)		(0.0001)
constant	4.1083	edu year	0.0028
	(0.0001)		(0.0001)
Unemployment value		constant	-2.2334
			(0.0001)
marriage	0.1068	Job separation rate	
	(0.0001)		
num of children	0.7443	edu year	0.0005
	(0.0001)		(0.0001)
age	0.0014	constant	-3.9982
	(0.0001)		(0.0001)
$age^2 \times 100$	-0.0354	Return Migration Cost	
	(0.0001)		
constant	0.0276	marriage	0.8663
	(0.0001)		(0.0001)
Migration cost		num of children	0.0798
			(0.0001)
social network	-1.6364	cohort	-0.0034
	(0.0001)		(0.0001)
marriage	0.1534	constant	8.7035
	(0.0001)		(0.0001)
num of children	0.0395		
	(0.0001)		
cohort	-0.1378		
	(0.0001)		
constant	15.5154		
	(0.0001)		

1. sn_{it} is an indicator variable that takes value 1 if the individual has social networks.
2. expu and expr stand for work experience in urban and rural areas respectively. They are both measured in months. Age is measured in years.
3. Cohort is defined by the birth year-1999.
4. Marriage is an indicator of marital status that takes value 1 if the individual is married.
6. Numbers in parentheses are standard errors.

Table C.3: The Impact of Social Networks: Based on the Model without Network Investment Decision

	Migration Cost	Job Arrival Rate
Social networks coefficient	-1.64	0.74
Average		
With Networks	91.64%	0.19
Without Network	100%	0.10

1. I normalize the average migration costs for rural migrants without social networks in the model to 1. Hence, migration costs are presented as relative to 1, (i.e., 89.19% means that the average migration costs for individuals with networks is 89.19% for those without social networks.)
2. The first panel in the job arrival rate column presents the point estimate for social networks in the job arrival rate equation. The second panel shows the calculated average job arrival rate for rural migrants who are not employed depending on their network status.

C.2 Estimation and Simulation Results

Table C.4: Model Fit: Earnings (Based on the Model without Network Investment Decision)

	Data	Model
migrants		
log(earnings)	7.1068	7.1447
$sd(log(earnings))$	0.2681	0.2504
with networks	7.1229	7.1467
$sd(log(earnings))$	0.2695	0.2497
without networks	7.0580	7.1365
$sd(log(earnings))$	0.2612	0.2528
non-migrants		
log(earning)	5.0310	5.0205
$sd(log(earnings))$	1.6292	1.6303
with networks	5.0677	5.0334
$sd(earnings)$	1.5418	1.6279
without networks	4.9410	4.9908
$sd(earnings)$	1.8318	1.6346

1. Migrants include people who currently work in urban areas.
2. sd(log(earnings)) stands for the standard deviation of log earnings.

Table C.5: Model Fits: Choices (Based on the Model without Network Investment Decision)

	Data	Without Investment Choice
social networks	72.30%	72.38%
*migrants**	29.00%	27.12%
with networks	21.86%	21.56%
without networks	7.14%	5.56%
return migrants	0.67%	1.06%
with networks	0.45%	0.80%
without networks	0.22%	0.26%
moving (rural to urban)	0.76%	1.18%
with networks	0.54%	0.97%
without networks	0.22%	0.22%
job search duration	1.91	2.34

1. The data column provides the moments calculated based on the observations during 2007-2009.
2. The numbers in the table are averages dividing the 36 months (2007-2009) observed in the data.

Table C.6: Counterfactual Results (Based on the Model without Network Investment Choice)

			Social Networks Only Affect	
	Model	Neither	Migration Cost	Job Arrival Rate
social networks	72.38%	72.38%	72.38%	72.38%
migrants	27.12%	17.22%	19.37%	26.82%
with networks	21.56%	12.06%	14.29%	20.76%
without networks	5.56%	5.16%	5.08%	6.06%
non-migrants	72.88%	82.78%	80.63%	73.18%
with networks	49.27%	60.31%	58.08%	51.61%
without networks	23.61%	22.47%	22.55%	21.43%
job search duration	2.34	2.68	2.55	2.51

1. Migrants include people who were born in rural areas and resided in urban cities who can be employed or unemployed.
2. The column of model shows the benchmark values simulated by the model estimates.
3. The neither column gives the simulation results when social networks affect neither migration costs nor the job arrival rate.
4. The column of migration cost gives the counterfactual results if social networks only reduce migration costs.
5. The column of job arrival rate gives the counterfactual results if social networks only increase the job arrival rate.
6. The row of equalization monthly tax shows how much individuals would like to pay per month for their lives on average to achieve the same utility when not allowing them to invest in their social networks.

Appendix D

Appendix for Chapter 3: Theorem

Semiparametric Identification

The following assumptions are used throughout this section in order to prove semiparametric identification of all the elements of the model. For simplicity, delete the i subscript. Let $U_s = (U_{s,1}, ..., U_{s,T})$, $U = (U_{hs}, U_{col})$, $\xi^H = \left(\xi_1^H, ..., \xi_T^H\right)$, $\xi^C = \left(\xi_1^C, ..., \xi_T^C\right)$, $\xi = \left(\xi^H, \xi^C\right)$.

(A-1) U, ζ and ξ *have distributions that are absolutely continuous with respect to Lebesgue measure with support* $Support(U) \times Support(\zeta) \times Support(\xi)$ *that may be bounded or infinite. Variances are assumed to be finite. The cumulative distribution function of* ζ *is assumed to be strictly increasing over its full support.*[1]

(A-2) $(X, K, Z) \perp\!\!\!\perp (U, \zeta, \xi)$ (Independence)

Semiparametric Identification of Log-wages with and without a Factor Structure

Identification of the log-wage equations is proved in theorem D.0.1. Only the case in which the measurements are continuous is considered, but, as shown in Carneiro, Hansen, and Heckman (2003), the measurements could also be discrete or mixed discrete-continuous.[2]

Theorem D.0.1 Let $\mu_s(X) = (\mu_{s,1}(X), ..., \mu_{s,T}(X))$. *Assume that the relevant elements of* ***(A-1)*** *and* ***(A-2)*** *(i.e., the joint conditions on* X, Z, U, ζ*) hold and that the following variation free condition holds:*

(A-3) $Support(\phi(Z), \mu_s(X)) = Support(\phi(Z)) \times Support(\mu_s(X))$. (Variation free)

Assume that $Support(\phi(Z)) \supseteq Support(\zeta)$ and $Support(\mu_s(X)) \supseteq Support(U_s)$. Then, the mean functions $\mu_{s,t}(X)$ are identified on the support of X. Also, the joint distribution of U_s is nonparametrically identified for $t = 1, ..., T$ for each $s = hs, col$.

[1]This assumption can easily be relaxed and is only made for convenience.

[2]In all cases, with additional assumptions, we can relax additive separability and identify functions of the form $y = \mu(X, U)$ by using the analysis in Matzkin (2003).

Proof Under the conditions of the theorem, we can find limit sets $\mathcal{Z}^{-}$and a $\mathcal{Z}^{+}$such that
$Pr(S = col \mid Z \in \mathcal{Z}^{-}) = 0$ and $Pr(S = col \mid Z \in \mathcal{Z}^{+}) = 1$ where we can still freely change the $\mu_s(X)$. Identification of the mean functions over their support is trivial since we observe $\ln W_s$ for each X and can recover the marginal distribution of U_s. The intercepts are recovered from assumed zero mean of U_s. The joint follows immediately since $Pr(\ln W_s < w \mid X) = F_{U_s}(w - \mu_s(X))$ by assumption **(A-2)**. Then, from **(A-3)** we can find an $X = x$ where $\mu_s(x) = k$ and k is a T dimensional vector. Let $m = k - \mu_s(x)$ so $Pr(\ln W_s < w \mid X = x) = F(k)$. Since the point w is arbitrary, we can vary it to identify the full joint distribution.

When the unobservables are represented in terms of equations (3.17), the next theorems show that we can nonparametrically identify the distributions of the factors and the uniquenesses as well as the factor loadings. We first state a theorem that will be useful for this purpose.

Theorem D.0.2 *Let Q_1 and Q_2 be two random variables that satisfy*

$$\begin{aligned} Q_1 &= \theta + R_1 \\ Q_2 &= \theta + R_2 \end{aligned}$$

where $\theta, R_1$and R_2 are mutually independent with $E(\theta) < \infty, E(R_1) = 0, E(R_2) = 0$, the conditions of Fubini's theorem are satisfied for each random variable and they have non-vanishing (a.e.) characteristic functions. Then, the marginal densities of θ, R_1 and R_2 are identified.

Proof See Kotlarski (1967), Prakasa Rao (1992).

Consider using only the information wages in each schooling state. For a given schooling level s we have a system of equations

$$\begin{aligned} lnW_{s,1} &= \mu_{s,1}(X_{s,1}) + \theta\alpha_{s,1} + \varepsilon_{s,1} \\ &\vdots \\ lnW_{s,T} &= \mu_{s,T}(X_{s,T}) + \theta\alpha_{s,T} + \varepsilon_{s,T} \end{aligned} \tag{D.1}$$

The total number of equations is given by T, while the total number of factors is given by $T - 1$ (an additional factor per-period up to $T - 1$). Such an arrangement would be motivated by the assumptions about the arrival of information made in the text. If we rearrange the equations in (D.1) and put the factor loadings in a matrix, it would have the triangular form

Time	Loadings for factor						
	θ_1	θ_2	θ_3	...	θ_{T-3}	θ_{T-2}	θ_{T-1}
1	$\alpha_{s,1,1}$	0	0	...	0	0	0
2	$\alpha_{s,2,1}$	$\alpha_{s,2,2}$	0	...	0	0	0
3	$\alpha_{s,3,1}$	$\alpha_{s,3,2}$	$\alpha_{s,3,3}$	...	0	0	0
...	...	...	...	...	...	...	...
$T-2$	$\alpha_{s,T-2,1}$	$\alpha_{s,T-2,2}$	$\alpha_{s,T-2,3}$	...	$\alpha_{s,T-2,T-3}$	$\alpha_{s,T-2,T-2}$	0
$T-1$	$\alpha_{s,T-1,1}$	$\alpha_{s,T-1,2}$	$\alpha_{s,T-1,3}$	...	$\alpha_{s,T-1,T-3}$	$\alpha_{s,T-1,T-2}$	$\alpha_{s,T-1,T-1}$
T	$\alpha_{s,T,1}$	$\alpha_{s,T,2}$	$\alpha_{s,T,3}$	...	$\alpha_{s,T,T-3}$	$\alpha_{s,T,T-2}$	$\alpha_{s,T,T-1}$

We first illustrate identification of the high school wages system of equations. For simplicity we illustrate it for the case in which the distribution of the factors is non-symmetric, but a (more elaborate) proof can be given for the case of symmetric distributions.

Theorem D.0.3 *From the analysis in Theorem D.0.1 we have data on $F(U_{hs} \mid X)$. Assume that U_{hs} has a factor structure representation as in* (3.17) *and that*

(A-4) $E\left(\theta_\ell^{k_\ell}\right) \neq 0$ *for $\ell = 1, ..., T-1$ and k_ℓ an odd integer.*

Then, the loadings $\{\alpha_{hs,t}\}_{t=1}^{T}$ are identified up to one normalization for each factor. The marginal distributions of $\left\{\theta_j\right\}_{j=1}^{T-1}$ and $\{\varepsilon_{hs,t}\}_{t=1}^{T}$ are nonparametrically identified as well.

Proof Notice that, since the factors have no natural scale, we need to set it (that is $\delta\theta = \kappa\delta\frac{\theta}{\kappa}$ for any constant κ). We also need to normalize the sign of the effect of the factor since, for example, having more of factor ℓ and $\alpha_{hs,t,\ell} > 0$ is equivalent to having less of the factor and $\alpha_{hs,t,\ell} < 0$. To pin down sign and scale, we normalize one loading to one for each factor.

Start by taking the first equation and an arbitrarily chosen ℓ^{th} equation. Without loss of generality, we normalize $\alpha_{hs,1,1} = 1$. Since we know the joint distribution of the unobservables in these equations we can identify the loadings on the first factor for the ℓ^{th} equation by forming

$$\frac{E\left(U_{hs,1}U_{hs,\ell}^{k_1}\right)}{E\left(U_{hs,1}^{2}U_{hs,\ell}^{k_1-1}\right)} = \alpha_{h,\ell,1}.$$

Since the choice of the ℓ^{th} equation is arbitrary we can identify all of the loadings for factor one. With the loadings on hand, we can the take equation 1 and the ℓ^{th} equation and form

$$U_{hs,1} = \theta_1 + \varepsilon_{h,1},$$

$$\frac{U_{hs,\ell}}{\alpha_{hs,\ell,1}} = \theta_1 + \sum_{j=2}^{\min\{\ell,T-1\}} \theta_j\frac{\alpha_{hs,\ell,j}}{\alpha_{hs,\ell,1}} + \frac{\varepsilon_{hs,\ell}}{\alpha_{hs,\ell,1}}.$$

Using Theorem D.0.2 we can nonparametrically identify the distributions of $\theta_1, \varepsilon_{hs,1}$ and $\sum_{j=2}^{\min\{\ell,T-1\}} \theta_j \frac{\alpha_{hs,\ell,j}}{\alpha_{hs,\ell,1}} + \frac{\varepsilon_{hs,\ell}}{\alpha_{hs,\ell,1}}$.

We now take equation 2 and some arbitrary equation $\ell > 2$. Normalizing $\alpha_{hs,2,2} = 1$, we can identify the loadings on factor 2 on the remaining equations by forming

$$\frac{E\left((U_{hs,2} - \theta_1\alpha_{hs,2,1})\,(U_{hs,\ell} - \theta_1\alpha_{hs,\ell,1})^{k_2}\right)}{E\left((U_{hs,2} - \theta_1\alpha_{hs,2,1})^2\,(U_{hs,\ell} - \theta_1\alpha_{hs,\ell,1})^{k_2-1}\right)} = \alpha_{hs,\ell,2}.$$

Since the choice of ℓ is arbitrary, we can identify all of the loadings on factor 2. By applying Theorem D.0.2 as before, we can identify the distributions of $\theta_2, \varepsilon_{hs,2}$ and $\sum_{j=3}^{\min\{\ell,T-1\}} \theta_j \frac{\alpha_{hs,\ell,j}}{\alpha_{hs,\ell,2}} + \frac{\varepsilon_{hs,\ell}}{\alpha_{hs,\ell,2}}$ nonparametrically.

By proceeding sequentially we can identify all of the loadings and nonparametric distributions of wages for schooling $s = hs$.

From Theorem D.0.3, we now have knowledge of the nonparametric distribution of the factors and uniquenesses as well as the loadings for wages for $s = h$. We next turn our attention to identification to the analogous system of equations in (D.1) for college ($s = col$).

Corollary D.0.4 *Under the assumptions of Theorem D.0.3, both the loadings $\alpha_{col,t}$ and the nonparametric distribution of $\varepsilon_{col,t}$, $t = 1,...T$, are identified without further normalizations.*

Proof To see why, take $U_{col,1}$ and an arbitrarily chosen period ℓ. Now form

$$\frac{E\left(U_{col,1}^2 U_{col,\ell}\right)}{E\left(U_{col,1} U_{col,\ell}\right)} = \alpha_{col,1,1}\frac{E\left(\theta_1^3\right)}{E\left(\theta_1^2\right)}.$$

Since we know the distribution of θ_1 from Theorem D.0.3, we know $\frac{E\left(\theta_1^3\right)}{E\left(\theta_1^2\right)}$ and hence we can recover $\alpha_{col,1,1}$. Now, by forming

$$cov\left(U_{col,1}U_{col,\ell}\right) = \alpha_{col,1,1}\alpha_{col,\ell,1}E\left(\theta_1^2\right)$$

we can recover $\alpha_{col,\ell,1}$. Since ℓ was chosen arbitrarily, we can recover the loadings for factor 1 for all periods. The nonparametric distribution of $\varepsilon_{col,1}$ can be recovered from $\frac{U_{col,1}}{\alpha_{col,1,1}} = \theta_1 + \frac{\varepsilon_{col,1}}{\alpha_{col,1,1}}$ by deconvolution. Proceeding sequentially the loadings on the remaining factors as well as the nonparametric distributions of the remaining uniquenesses can also be recovered.

Identification of the Cost Function

Write the college attendance condition (3.8) as $E(\mathcal{V}_{col,1}(\mathcal{I}_1) - \mathcal{V}_{hs,1}(\mathcal{I}_1) - \phi(Z) - \theta\lambda - \omega \mid \mathcal{I}_0) > 0$. Let θ^0 denote the elements of θ included in the agent's information set at the time the schooling decision is made, and let λ^0 denote the sub-vector of λ

associated with them. Define $E\left(\mathscr{V}^*_{col,1}(\mathcal{I}_1) - \mathscr{V}^*_{hs,1}(\mathcal{I}_1) \mid \mathcal{I}_0\right) = \mu_V(X) + \tau\left(X, \theta^0\right)$ which is known since all of the elements are known from our previous analysis. The econometrician has data on the left hand side of

$$\Pr(S = col|X, Z) = \Pr\left(\tau\left(X, \theta^0\right) - \theta^0\lambda^0 - \omega > \phi(Z) - \mu_V(X)\right).$$

Theorem D.0.5 *Assume that the relevant elements of* **(A-1)** *hold. Change* **(A-2)** *so that the independence of Z and the error terms holds conditional on X. Let Z^e be the elements of Z that are not a part of X (excluded from wages) and further assume that we can define $\phi(z^e, x)$ for all pairs (z^e, x) in the support of Z. As with all discrete choice problems the scale needs to be set. Assume that*

(A-5) $var\left(\tau\left(\widetilde{x}, \theta^0\right) - \theta^0\lambda^0 - \omega\right) = 1$ *for* $X = \widetilde{x}$.

Then, if $\phi(Z)$ satisfies the is part of the Matzkin class of functions ((Matzkin, 1992; Heckman and Navarro, 2007), λ^0 and the nonparametric distribution of ω are identified up to normalization.

Proof Define $\Upsilon\left(X, \theta^0\right) = \tau\left(X, \theta^0\right) - \theta^0\lambda^0 - \omega$ and fix $X = \widetilde{x}$. The observed probability that the agent chooses college (conditional on $X = \widetilde{x}$ and using **A-3**) is $\Pr\left(\Upsilon\left(\widetilde{x}, \theta^0\right) > \phi(Z^e, \widetilde{x}) - \mu_V(\widetilde{x})\right)$ where $\mu_V(\widetilde{x})$ is a known constant (conditional on $X = \widetilde{x}$). Using the (conditional on X) independence of $\theta^0\lambda_0 + \omega$ and Z^e we can then use the analysis of Matzkin (1992) to identify $\phi(Z^e, \widetilde{x})$ and the distribution of $\Upsilon\left(\widetilde{x}, \theta^0\right)$, all of this conditional on $X = \widetilde{x}$, up to a normalization.

Next, still conditional on $X = \widetilde{x}$, we can form the joint distribution of log wages and the choice index since we know the left hand side of

$$\Pr\left(U_s \leq \ln W_s - \mu_s(\widetilde{x}) \mid X = \widetilde{x}, S = s, Z^e = z^e\right)\Pr\left(S = s \mid Z^e = z^e\right)$$

$$= \int_{-\infty}^{\ln W_s - \mu_s(\widetilde{x})} \int_{\phi(z^e, \widetilde{x}) - \mu_V(\widetilde{x})}^{\infty} f\left(U_s, \Upsilon\left(\widetilde{x}, \theta^0\right)\right) d\Upsilon\left(\widetilde{x}, \theta^0\right) dU_s.$$

We can trace this integral by varying $\ln W_s, z_E$, so we can identify the joint distribution of $U_s, \Upsilon\left(X, \theta^0\right)$. With the joint distribution in hand, we can follow the reasoning of theorem D.0.1 to identify λ^0. To see how, take the covariance between the choice index and the first equation for high school :

$$cov\left(\Upsilon\left(\widetilde{x}, \theta^0\right), U_{h,1}\right) = cov\left(\tau\left(\widetilde{x}, \theta^0\right), \theta_1\right) + \lambda_1^0\sigma_{\theta_1}^2$$

where everything but λ_1^0 is known so we can solve for it. Proceeding recursively we identify all of the elements of λ^0.

With λ^0 known, we can take $\Upsilon\left(\widetilde{x}, \theta^0\right)$ and deconvolve the distribution of ω since it is independent of X, θ by **A-2** and the factor structure assumptions. Finally, by changing the value of $z^e, \widetilde{x}$ we can trace $\phi(Z^e, X)$ coordinate by coordinate since, for any value of X, the scale $var\left(\tau\left(X, \theta^0\right) - \theta^0\lambda_0 - \omega\right)$ is a function of known elements.

Appendix E

Appendix for Chapter 3: Data

Data

We use data on white males from NLSY79 and pool it with data for white males that are household heads from PSID. In the original NLSY79 sample there are 2439 white males. Out of this, 1334 have either a high school degree (and high school only) or a college degree. We then try to recover earnings for as many individuals as possible. First, individual earnings are formed by taking total income from wages and salary in the past calendar year directly from the NLSY deflated to year 2000 prices using the Consumer Price Index reported by the Bureau of Labor Statistics. Then, because the NLSY survey was not administered on even years starting in 1995, earnings for any individual in these years are not observed. We impute earnings in these years by taking the average of the earnings in the immediately adjacent years if available. Unemployed individuals have zero earnings on those years and are flagged as having missing earnings otherwise.

The PSID sample is a little more problematic since attrition is more common in that survey than in the NLSY79.[1] While this reduces the PSID sample size available to us, it allows us to analyze people of all ages, something that cannot be done with the NLSY79. Earnings in the PSID sample are obtained by using the annual labor earnings variable. As with NLSY79, we impute earnings for the years in which there was no survey using an average of the two immediately adjacent years if possible.

The individual life (ages 18 to 65) is then simplified into 10 periods: $t = 0$ (schooling choice decision, right before 18), $t = 1$ (18-23), $t = 2$ (24-29), $t = 3$ (30-35), $t = 4$ (36-41), $t = 5$ (42-47), $t = 6$ (48-53), $t = 7$ (54-59), $t = 8$ (60-65) and $t = 9$ for age above 65. Earnings for each period are defined as the present value of earnings for the ages included in the period discounted using an interest rate of 3%. If the present value of earnings cannot be formed then the individual is flagged as having miss-

[1]But see Fitzgerald, Gottschalk, and Moffitt (1998) for evidence that observed people in PSID have similar characteristics as those in the CPS so attrition is roughly random.

ing present value of earnings for the period. Total working hours are then formed by adding total hours worked during the period. Finally, we form "wages" for the period by dividing present value of earnings over hours worked.

The procedure to get consumption data in both samples is fairly similar in principle. Household consumption at time period t is defined as the difference between available resources -household income plus assets available at the beginning of the period- minus the discounted assets available the next period.[2] In order to recover more consumption observations, if assets are not observed at the beginning of the period as required, we use assets either one year before or after it (discounted appropriately). Imputing household consumption for the PSID sample is done in the same way as with NLSY79, although questions about assets are not asked as frequently for this sample. To impute individual consumption we divide household consumption by the square root of the number of members in the household. We try to correct for the measurement error introduced by this procedure in the estimation as explained in the text.

Tuition between 1972 and 2000 is defined as the average in state tuition in colleges in the county of residence. If there is no college in the county then average tuition in the state is taken instead. For years prior to 1972, national tuition trends are used to project county tuition backwards keeping the average observed structure of the period 1972-1977. That is, a regression of the difference between national tuition and county tuition between 1972 and 1977 against county dummies is run. The predicted value is the average difference between national and county tuition which can then be added to the national tuition observed in the years previous to 1972.

Ability specific variables

In 1980, NLSY respondents were administered a battery of ten achievement tests referred to as the Armed Forces Vocational Aptitude Battery (ASVAB) (See Cawley, Conneely, Heckman, and Vytlacil (1997) for a complete description). The math and verbal components of the ASVAB can be aggregated into the Armed Forces Qualification Test (AFQT) scores.[3] Many studies have used the overall AFQT score as a control variable, arguing that this is a measure of scholastic ability. In this paper, the interpretation that AFQT is an imperfect proxy for scholastic ability is taken and the factor structure is used to capture this. Potential aggregation bias is avoided by using each of the components of the ASVAB score as a separate measure. PSID participants do not take the ASVAB battery of tests that NLSY participants do. Instead, we use the IQ Word test that was administered in 1972 and assume

[2]See Browning and Leth-Petersen (2003) for evidence that, at least for the case of Denmark where very detailed data on consumption is available, this procedure works well.

[3]Implemented in 1950, the AFQT score is used by the U.S. Army to screen draftees.

that this test (measured in percentiles) is comparable to the ASVAB tests (also in percentiles).

Appendix F

Appendix for Chapter 3: Tables

Tables

Table A3-1

Ability Measurement System

	NLSY					PSID
	Reasoning	**Word Knowledge**	**Paragraph Composition**	**Coding Speed**	**Math Knowledge**	**IQ Word Test**
Age at Test Date	-0.020	-0.007	-0.025	-0.020	-0.038	0.004
	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)	(0.000)
Enrolled at Test Date	0.147	0.165	0.145	0.079	0.205	0.076
	(0.020)	(0.019)	(0.020)	(0.017)	(0.019)	(0.014)
Grade Completed at Test Date	0.081	0.085	0.085	0.075	0.101	0.018
	(0.010)	(0.009)	(0.010)	(0.008)	(0.009)	(0.001)
Ability	1.000	0.764	0.800	0.598	0.902	1.111
	-	(0.030)	(0.034)	(0.033)	(0.024)	(0.043)
Constant	-0.088	-0.390	-0.006	0.009	0.002	0.196
	(0.097)	(0.090)	(0.097)	(0.083)	(0.088)	(0.023)
σ_ε^2	0.015	0.035	0.042	0.052	0.015	0.001
	(1.000)	(1.000)	(1.001)	(1.001)	(1.000)	(1.000)

1. Standard Errors in parenthesis.

2. Measurement j is given by $M_{i,j} = X_i\beta_j^M + Q_i\alpha_j^M + \varepsilon_{i,j}^M$.

Table A3-2

Distribution of Ability

Mixture Component	1	2	**3**
Mean	0.144	-0.171	-0.219
	(0.014)	(0.026)	-(0.024)
Variance	0.034	0.042	0.038
	(0.003)	(0.006)	(0.007)
Weights	0.575	0.210	0.215
	(0.350)	(0.210)	(0.440)

1. Standard Errors in parenthesis.

2. Measurement j is given by $M_{i,j} = X_i\beta_j^M + Q_i\alpha_j^M + \varepsilon_{i,j}^M$, where Q_i follows a mixture of normals distribution with 3 components.

Table A3-3

Log Wages Coefficients

High School Graduates

	Ability	**Age**	**Age^2**	**NLSY**	**Constant**
	0.080	0.357	-0.004	-1.069	-40.600
	(0.144)	(0.002)	(0.000)	(0.117)	(0.110)

Factor Loadings:	**θ1**	**θ2**	**θ3**	**θ4**	**θ5**	**θ6**	**θ7**	**$\sigma^2_{\varepsilon,hs,t}$**
Age 18-23	1.000							1.923
	-							(0.248)
Age 24-29	1.867	1.000						1.874
	(0.086)	-						(0.166)
Age 30-36	1.916	1.576	1.000					0.666
	(0.081)	(0.165)	-					(0.126)
Age 36-41	1.807	1.450	1.019	1.000				1.208
	(0.084)	(0.170)	(0.143)	-				(0.118)
Age 42-47	1.714	1.980	0.477	2.016	1.000			1.701
	(0.111)	(0.157)	(0.174)	(0.148)	-			(0.132)
Age 48-53	1.854	2.322	-0.333	2.059	1.439	1.000		1.771
	(0.112)	(0.222)	(0.203)	(0.232)	(0.157)	-		(0.209)
Age 54-59	1.436	2.140	-1.178	2.922	2.002	1.924	1.000	0.091
	(0.095)	(0.137)	(0.101)	(0.106)	(0.072)	(0.102)	-	(0.028)
Age 60-65	1.531	4.178	-2.045	0.119	1.273	3.352	0.596	0.364
	(0.156)	(0.215)	(0.226)	(0.126)	(0.084)	(0.146)	(0.246)	(0.065)

College Graduates

	Ability	**Age**	**Age^2**	**NLSY**	**Constant**
	1.917	0.838	-0.010	-0.737	-48.520
	(0.096)	(0.000)	(0.000)	(0.041)	(0.051)

Factor Loadings:	**θ1**	**θ2**	**θ3**	**θ4**	**θ5**	**θ6**	**θ7**	**$\sigma^2_{\varepsilon,col,t}$**
Age 18-23	-0.268							3.526
	(0.188)							(0.341)
Age 24-29	-1.833	0.957						1.294
	(0.056)	(0.148)						(0.111)
Age 30-36	-1.470	1.057	1.425					1.063
	(0.064)	(0.132)	(0.086)					(0.119)
Age 36-41	-1.472	0.003	2.686	-0.107				0.832
	(0.065)	(0.136)	(0.016)	(0.120)				(0.048)
Age 42-47	-1.563	-0.782	3.010	0.709	0.110			2.306
	(0.080)	(0.059)	(0.043)	(0.152)	(0.086)			(0.189)
Age 48-53	-1.645	-1.466	2.582	1.789	-0.190	0.267		3.939
	(0.054)	(0.012)	(0.030)	(0.007)	(0.026)	(0.015)		(0.078)
Age 54-59	-2.519	-2.205	2.556	2.681	-1.268	-0.243	1.405	0.371
	(0.010)	(0.027)	(0.016)	(0.028)	(0.030)	(0.007)	(0.034)	(0.009)
Age 60-65	-3.482	-3.199	0.079	2.081	-2.810	-1.027	1.112	1.480
	(0.010)	(0.027)	(0.016)	(0.028)	(0.030)	(0.007)	(0.034)	(0.081)

1. Standard Errors in parenthesis.

2. Logwages in schooling s at time t are given by $\ln W_{i,s,t}=X_{i,t}\beta_s+\theta_i\alpha_{s,t}+\varepsilon_{i,s,t}$, where each factor $\theta_{i,j}$ is distributed as a mixture of normals with 2 components.

Table A3-4

Factor Distribution

Mixture Component	1		2	
	Mean	Weights	Mean	Weights
θ_1	-1.114	0.454	0.927	0.546
	(0.003)	(0.001)	-	-
θ_2	0.225	0.394	-0.146	0.606
	(0.001)	(0.005)	-	-
θ_3	-0.371	0.590	0.535	0.410
	(0.007)	(0.003)	-	-
θ_4	-0.029	0.849	0.164	0.151
	(0.000)	(0.001)	-	-
θ_5	-0.683	0.558	0.863	0.442
	(0.006)	(0.002)	-	-
θ_6	0.352	0.159	-0.066	0.841
	(0.007)	(0.010)	-	-
θ_7	-0.270	0.418	0.194	0.582
	(0.013)	(0.010)	-	-

1. Standard Errors in parenthesis.

2. Logwages in schooling *s* at time *t* are given by $\ln W_{i,s,t}=X_{i,t}\beta_s+\theta_i\alpha_{s,t}+\varepsilon_{i,s,t}$, where each factor $\theta_{i,j}$ is distributed as a mixture of normals with 2 components, where the variance of each component is normalized to 0.5.

Table A3-5

Preference Parameters

Relative Risk Aversion Coefficient (ψ)	0.491	Discount Factor (1/(1+ ρ))	0.941
	(0.000)		(0.000)
Terminal Period		Labor	
Curvature Coefficient (χ)	0.088	Frisch Elasticity (φ)	0.996
	(0.000)		(0.004)
Weight (*b*)	0.196	Weight (*h*)	16.719
	(0.000)		(0.016)

1. Standard Errors in parenthesis.

2. Preferences are given by $\left[\frac{C^{1-\psi}}{1-\psi} - h\frac{n^{1+1/\varphi}}{1+1/\varphi}\right]$, and by $\left[b\frac{(Y_{\min}+(1+r)A_T)^{1-\chi}}{1-\chi}\right]$ in the terminal period.

Table A3-6

Measurement errors

	Consumption	Labor Supply
Age	0.031	-0.056
	(0.003)	(0.004)
Married	0.129	
	(0.043)	
Number of Children	-0.609	
	(0.014)	
NLSY	1.633	-2.778
	(0.083)	(0.187)
Age x NLSY		0.047
		(0.004)
Constant	-1.858	3.068
	(0.163)	(0.178)
Variance		
Age 18-23		0.538
		(0.035)
Age 24-29	0.490	1.568
	(0.018)	(0.031)
Age30-35	0.583	0.178
	(0.027)	(0.010)
Age36-41	0.819	0.248
	(0.051)	(0.016)
Age 42-47	0.904	0.393
	(0.058)	(0.022)
Age 48-53	1.006	2.598
	(0.148)	(0.064)
Age 54-59	0.524	0.552
	(0.121)	(0.053)
Age 60-65	0.625	1.405
	(0.186)	(0.118)

1. Standard Errors in parenthesis.

Table A3-7

College Education Psyhic Cost Coefficients

Ability	-10.320	NLSY	5.817
	(3.720)		(1.785)
Local Unemployment Rate for College Drop	4.673	Constant	12.477
	(27.150)		(3.924)
Local Unemployment Rate for College Grad	8.667		
	(40.476)		
Number of Siblings	2.146	θ_1	-7.131
	(0.430)		(2.165)
Mother's Education Year	0.111	θ_2	-12.154
	(1.563)		(1.329)
Mother's Education Year^2	-0.581	θ_7	22.007
	(0.185)		(3.145)

1. Standard Errors in parenthesis.

Table A3-8

Model Fit: Average Outcomes Conditional on Schooling Choice

	High School Graduates				College Graduates			
	Log Hours		Log Consumption		Log Hours		Log Consumption	
	Data	Predicted	Data	Predicted	Data	Predicted	Data	Predicted
Age 18-23	8.912	9.216			8.416	8.102		
		[9.023, 9.350]				[7.848, 8.277]		
Age 24-29	9.440	9.383	2.627	2.642	9.347	9.434	3.031	3.006
		[9.213, 9.532]		[2.532, 2.774]		[9.268, 9.599]		[2.855, 3.130]
Age 30-36	9.496	9.395	2.653	2.631	9.531	9.558	3.072	3.108
		[9.311, 9.461]		[2.516, 2.746]		[9.493, 9.619]		[2.966, 3.220]
Age 36-41	9.523	9.379	2.676	2.635	9.571	9.671	3.105	3.120
		[9.313, 9.458]		[2.481, 2.775]		[9.580, 9.734]		[2.971, 3.294]
Age 42-47	9.461	9.365	2.820	2.868	9.524	9.680	3.376	3.350
		[9.271, 9.447]		[2.687, 3.033]		[9.578, 9.770]		[3.163, 3.543]
Age 48-53	9.425	9.313	3.081	3.083	9.480	9.587	3.540	3.606
		[9.033, 9.646]		[2.700, 3.459]		[9.279, 9.901]		[3.312, 4.073]
Age 54-59	9.354	9.229	3.167	3.168	9.388	9.484	3.386	3.335
		[8.979,9.474]		[2.561, 3.645]		[9.188, 9.734]		[2.805, 3.961]
Age 60-65	8.771	8.755	2.997	3.347	8.948	8.995	3.612	3.527
		[8.398, 9.137]		[2.236, 4.333]		[8.612, 9.371]		[2.492, 4.510]

1. 95% Confidence interval in brackets.

Table A3-9

Model Fit: Average Log Wages conditional on Schooling Choice

	High School Graduates		College Graduates	
	Data	Predicted	Data	Predicted
Age 18-23	-7.250	-7.205	-7.361	-7.380
		[-7.280, -7.144]		[-7.471, -7.310]
Age 24-29	-6.819	-6.851	-6.624	-6.667
		[-6.916, -6.795]		[-6.749, -6.607]
Age 30-36	-6.653	-6.673	-6.322	-6.319
		[-6.738, -6.620]		[-6.393, -6.263]
Age 36-41	-6.532	-6.574	-6.042	-6.058
		[-6.637, -6.521]		[-6.145, -5.989]
Age 42-47	-6.529	-6.541	-5.931	-5.931
		[-6.609, -6.478]		[-6.025, -5.858]
Age 48-53	-6.432	-6.482	-5.923	-5.909
		[-6.579, -6.387]		[-6.041, -5.815]
Age 54-59	-6.444	-6.463	-5.957	-5.954
		[-6.618, -6.321]		[-6.183, -5.760]
Age 60-65	-6.626	-6.601	-6.049	-6.081
		[-6.810, -6.394]		[-6.372, -5.828]

1. 95% Confidence interval in brackets.

Curriculum Vitae

Name:	Jin Zhou
Post-Secondary Education and Degrees:	The University of Western Ontario Ph.D. in Economics 2010 - 2016 (Expected)
	Beijing Normal University Beijing, China 2007 - 2010 M.A. in Economics
	Beijing Normal University Beijing, China 2003 - 2007 B.A. in Finance
Honours and Awards:	Sir Arthur Currie Memorial Scholarship 2013 Western Graduate Research Scholarship 2010-2014 China Development Research Scholarship 2009-2010 Excellent Graduate Student in Beijing 2007
Related Work Experience:	Teaching Assistant The University of Western Ontario 2010 - 2014
	Research Assistant The University of Western Ontario 2011 - 2016