ESSAYS ON EDUCATIONAL ATTAINMENT

A Dissertation

by

YINGNING WANG

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

DOCTOR OF PHILOSOPHY

August 2009

Major Subject: Economics

ESSAYS ON EDUCATIONAL ATTAINMENT

A Dissertation

by

YINGNING WANG

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

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Li Gan
Committee Members,	Steven Puller
	Adalbert Mayer
	Joanna Lahey
Head of Department,	Larry Oliver

August 2009

Major Subject: Economics

ABSTRACT

Essays on Educational Attainment. (August 2009)

Yingning Wang, B.A., Central South University; M.A., Fudan University

Chair of Advisory Committee: Dr. Li Gan

One of the very interesting demographic features in the US over the last several decades is the persistent racial educational gap between blacks and whites and the reverse gender education gap as a result of the rapid rise in women's educational attainment. This dissertation is to investigate the reasons behind it.

I first investigated the educational gap between blacks and whites. I propose a new model to identify if and how much the educational attainment gap between blacks and whites is due to the difference in their neighborhoods. In this model, individuals belong to two unobserved types: the endogenous type who may move in response to the neighborhood effect on their education; or the exogenous type who may move for reasons unrelated to education. The Heckman sample selection model becomes a special case of the current model in which the probability of one type of individuals is zero. Although I cannot find any significant neighborhood effect in the usual Heckman sample selection model, I do find heterogeneous effects in our type-consistent model. In particular, there is a substantial neighborhood effect for the movers who belong to the endogenous type. No significant effects exist for other groups. On average, I find that the

neighborhood variable, the percentage of high school graduates in the neighborhood, accounts for about 37.7% of the education gap between blacks and whites.

This dissertation sheds some insight about women's educational attainment by studying the motivations of education for women: to pursue higher wages and to find highly educated spouses. The identification strategy is that the college education is exogenous to the partner choice if education is driven by pursuing higher job market return (the type of marry-for-romance), and is endogenous if the education decision is driven by marriage market return (the type of marry-for-money). I find that the marry-for-romance type has higher education than the other type and given everything else the same, with the same education level, the women who marry for money have a higher probability of finding a highly educated husband than those marrying for romance. Therefore, the reversal educational gap could be the result of more marry-for-romance women.

TABLE OF CONTENTS

Page

ABSTRACT iii

TABLE OF CONTENTS v

CHAPTER

I INTRODUCTION 1

II RESIDENTIAL MOBILITY, NEIGHBORHOOD EFFECTS, AND EDUCATIONAL ATTAINMENT OF BLACKS AND WHITES 5

2.1 Introduction 5
2.2 Data and descriptive statistics 11
2.3 The model 18
2.4 Estimation results 30
2.5 Conclusions 38

III MARRY FOR MONEY OR MARRY FOR ROMANCE: AN EMPIRICAL STUDY OF FEMALES' CHOICES OF EDUCATION AND PARTNERS 40

3.1 Introduction 40
3.2 Literature review 42
3.3 Empirical models 46
3.4 Data 51
3.5 Results 54
3.6 Conclusion 60

IV CONCLUSION 62

REFERENCES 65

APPENDIX A 73

Page

APPENDIX B 90

APPENDIX C 92

APPENDIX D 95

APPENDIX E...99

VITA 102

CHAPTER I

INTRODUCTION

One of very interesting demographic features in the US over the last several decades is persistent racial educational gap between blacks and whites and reverse gender education gap as a result of rapid rise in women's educational attainment. This dissertation is to investigate the reasons behind it.

The education gap between blacks and whites is substantial in the United States. According to the 2000 census, the percentage of whites graduating from high school is 18% more than that of blacks, while the percentage of whites with bachelor degrees is twice as much as that of blacks. A growing literature claims that neighborhoods and peer groups can be important in determining education outcomes (Durlauf, 2002; Manski, 1993). However, it remains difficult to empirically identify the neighborhood effect on education because individuals may move in response to the impact of the neighborhood on their education outcomes. Although this type of endogenous moving may have significant effects, it is only one of many reasons people may move. Individuals may choose to move for several other reasons which are independent of the neighborhood effect on education.

This dissertation follows the style of *The Journal of Human Resources*.

Chapter II argues that heterogeneous moving actually provides an opportunity to reveal individuals' preferences about neighborhoods and to help identify the neighborhood effect. In particular, I suggest that there are two unobserved types of individuals.For one type, moving is endogenous to the education decision, while for another type, moving is exogenous. Although the types are unobserved, it is possible to assign a probability that each household belongs to one of the two types. The usual Heckman sample selection model becomes a special case of this model in which the probability of falling into one of the two types equals zero or one.

A specification test rejects the Heckman model and favors our type-consistent model. Although I cannot find any significant neighborhood effect in the usual Heckman sample selection model, I do find heterogeneous effects in our type-consistent model. In particular, there is a substantial neighborhood effect for movers who belong to the endogenous type. No significant effects exist for other groups. On average, I find that the neighborhood effect accounts for about 37.7% of the education gap between blacks and whites.

Another key demographic feature in the US over the last several decades is the rapid rise in women's educational attainment. The educational gap between females and males has been declined or even reversed. More than half of American college students are women and about 60% of college graduates are females (Becker and Posner, 2008).

Chapter III tries to shed some insight by studying the motivations of education for women. An obvious return for more schooling is a higher wage in the labor market. In addition, the literature has provided another potential return for education: a better educated person may be able to meet a better educated partner in the marriage market. While the benefits from the marriage market apply for both genders, it is likely more important for women since married man on the average have higher labor force participation rates and higher income than married women.

There is a wide literature base about the effect of education on the quality of partners and it finds positive assortive mating in terms of education (Boulier 1984; Mare 1991; Cancian et al. 1993; Behrman et al. 1994; Juhn and Murphy 1997; Weiss and Willis 1997). Little work, however, investigates whether the mating and the education investment will be different if the education decisions are driven by different incentives. Chapter III focuses on empirically identifying the motivation of education and its effect on the prospect of partners. Chapter III first suggests an identification strategy to distinguish these two types of motivations for education. In particular, if the woman's education decision is endogenous to the partner choice, then pursuing marriage market return is her motivation for education(i.e., the woman is the "marry-for-money" type). However, if the woman's education decision is exogenous to the partner choice, then her education decision is driven by the labor market return (i.e., the woman is the "marry-for-romance" type). Second, I propose and estimate an empirical model that incorporates these two types of motivations in the market. The central empirical challenge is how to identify the type, since the type is unobservable. Chapter III suggests modeling the unobserved types by a mixture density to characterize the unobserved types. The estimation results suggest that both motivations exist in the population. A woman would have a probability of 0.388 to be the "marry-for-money" type, and the "marry-for-money" type would take 0.855 years less education than the "marry-for-romance" type. Further, at the same education level, the "marry-for-money" type woman would be much more likely to have a better educated partner than the "marry-for-romance" type. Finally, I find that a woman's attitude toward a female's role in a family is the key factor that identifies the two types of women. A woman with a more traditional view on the female's role in a family would be more likely to belong to the "marry-for-money" type.

This result suggests a potential reason for a recent rapid rise in women's educational attainment. The changes of women's attitude about the women's status in the family and in the society are documented in the literature. With more and more women becoming career oriented, they are more attracted by the financial benefit from the higher education. Therefore marriage consideration becomes less and less important in their education decision. The more marry-for-romance type women the society has, the higher the educational attainment the women will have.

CHAPTER II

RESIDENTIAL MOBILITY, NEIGHBORHOOD EFFECTS, AND EDUCATIONAL ATTAINMENT OF BLACKS AND WHITES

2.1 Introduction

The education gap between blacks and whites is substantial in the United States. According to the 2000 census, the percentage of whites graduating from high school is 18% more than that of blacks, while the percentage of whites with bachelor degrees is twice as much as that of blacks. A growing literature claims that neighborhoods and peer groups can be important in determining education outcomes (Durlauf, 2002; Manski, 1993). However, it remains difficult to empirically identify the neighborhood effect on education because individuals may move in response to the impact of the neighborhood on their education outcomes. Although this type of endogenous moving may have significant effects, it is only one of many reasons people may move. Individuals may choose to move for several other reasons which are independent of the neighborhood effect on education.

This chapter argues that heterogeneous moving actually provides an opportunity to reveal individuals' preferences about neighborhoods and to help identify the neighborhood effect. In particular, I suggest that there are two unobserved types of individuals. For one type, moving is endogenous to the education decision, while for another type, moving is exogenous. Although the types are unobserved, it is possible to assign a probability that each household belongs to one of the two types. The usual

Heckman sample selection model becomes a special case of this model in which the probability of falling into one of the two types equals zero or one.

A specification test rejects the Heckman model and favors our type-consistent model. Although I cannot find any significant neighborhood effect in the usual Heckman sample selection model, I do find heterogeneous effects in our type-consistent model. In particular, there is a substantial neighborhood effect for movers who belong to the endogenous type. No significant effects exist for other groups. On average, I find that the neighborhood effect accounts for about 37.7% of the education gap between blacks and whites.

The difference in returns to education between blacks and whites is often considered to be one of the reasons for the education gap. However, Neal and Johnson (1996) and O'Neil (1990) argue that the returns to education between blacks and whites have been converging over the last twenty years, while the education disparity has remained the same (Couch and Daly, 2002).

While the returns to education between blacks and whites have been converging over the past twenty years, the widespread residential segregation of blacks from whites in metropolitan areas remains (Cutler, Glasear and Vigdor, 1999). Segregation has declined in the cities but has increased in suburbia, resulting in little net change (Weinberg and Iceland, 2002). Wide racial segregation indicates that blacks and whites have different neighborhoods (Ananat, 2007). For example, the percentage of high school graduates in the black neighborhoods is only 59.9% of that in the white neighborhoods (see table A1.1-2).

Neighborhood effects on educational attainment have been widely studied. Earlier studies such as the Coleman Report (Coleman et al., 1979) find that the socioeconomic composition of students has significant effects on the unequal academic attainments of white and black children. Crane's (1991) test supports the existence of neighborhood effects by examining the pattern of neighborhood effects on dropout rates and teenage childbearing. Aaronson (1998) also finds neighborhood effects on high school graduation. However, other scholars find mixed evidence or an insignificant effect of neighborhoods on educational attainment (Brooks-Gunn, et al 1997; Datcher, 1982; Duncan, 1994; Duncan, et al. 1997; Lillard, 1993; Plotnick and Hoffman, 1999). The neighborhood variables in these studies include economic conditions, occupational composition, racial characteristics, poverty status, and demographic composition.

According to Ginther, Haveman and Wolfe (2000), the inconsistency of neighborhood effects on education actually implies that the research of neighborhood effects could be subject to omitted variable or selection biases. The difficulty in identifying neighborhood effects is well known within the literature and has thus received extensive attention. This difficulty arises because non-random self-selection generates the correlation between individual and group attributes, some of which are likely to be unobservable, resulting in biases in the estimation of group effects on group members' behaviors.

One approach to solve this problem is to randomize neighborhood choice by special social experiments. However, results from experiments are also mixed. Early results from the Gautreaux program imply that outcomes for parents and children were markedly better for those who moved to the less-segregated suburbs (Popkin, Rosenbaum and Meaden 1993). On the other hand, Kling and Votruba (2001) find that the placement assignments in the Gautreaux program were not entirely random. Moving to Opportunity (MTO) uses a randomized design and finds that randomly selected families who move from deprived areas to more affluent neighborhoods have less violent criminal behavior by teens, better health care and child care, and better mental and physical health. However, the difference between control groups and treatment groups in welfare participation, employment, and child test scores is much less than what is found in the Gautreaux studies (Hanratty, McLanahan and Pettit 1998; Katz, Kling and Liebman 2001; Ladd and Ludwig, 1997; Ludwig, Ladd, and Duncan, 2001). Similarly, Oreopolous (2003) and Jacob (2004) study the impact of public housing projects in Toronto and Chicago and find no significant differences in test scores and dropout rates. Finally Gibbons (2002) finds that educational attainment is slightly higher for those from neighborhoods with above average educated households by using data from contrasting council tenant housing in the United Kingdom. However, his result is not robust to random housing assignment.

The above results from social experiments suggest that the significance of neighborhood effects on education is sensitive to whether or not projects have random designs. Significant neighborhood effects on education are only found in experiments where housing assignment is not totally random. This implies that whether the residential mobility is endogenously chosen or not is strongly related with the extent of neighborhood effects on education.

Social experiments, although very insightful, are limited in scope. Another broad approach is to find econometric solutions (Brock and Durlauf, 2001; Manski, 1993; Moffitt, 2001). This approach concentrates on the endogeneity of the moving decisions and the neighborhood characteristics. Most recently, Ioannides and Zabel (2008) suggest an identification strategy by taking housing demand and neighborhood choice as joint decisions in the presence of neighborhood effects.

This chapter contributes to the econometric solution to the neighborhood effect problem by focusing on the endogenous selection process. I first show the heterogeneity in selection process for movers and non-movers. The coefficients for the inverse mills ratios in the Heckman sample selection model are significant for non-movers but not significant for movers. Therefore, I propose a new model to account for heterogeneity in the selection process. In this new model, individuals are assumed to belong to one of the two types: an endogenous type whose moving decisions and education decisions are jointly determined, and an exogenous type whose decisions about moving and education are determined independently. Because the type is unobserved, it can only be determined with a probability. Following the typical approach in the literature to model the unobserved types, I use a mixture density to characterize the unobserved types and the corresponding equations for schooling years.

Another difficulty in modeling residential mobility is that I only observe individuals' final residences but no alternative choices. Ignoring this problem will lead to inconsistent estimation. I solve this problem by choosing the average of characteristics within the group of counties that are spatially adjacent to the county where individuals reside as instrumental variables (Ioannides and Zanella, 2007).

Empirically, this chapter has two main findings. First, the chapter provides strong evidences of two unobserved types. The estimation in the chapter is conducted allowing the possibilities of either one type or two types. Clearly the two type model matches with the data better than the one type model. In addition, this chapter finds that people from different types do behave differently. The endogenous type would have taken 5.515 years of schooling, much more than 3.687 years of schooling taken by the exogenous type. Overall, 82.12% of people belong to the endogenous type.

Second, this chapter finds a strong neighborhood effect for the people who are endogenous and who are movers, despite no statistically significant effects are found in the one-type Heckman sample selection model. This result suggests the importance of type consistent model. In particular, the chapter finds that if a black who belongs to the endogenous type moves to a white neighborhood, his education would increase by 1.0965 years because of the neighborhood effect. Overall, the difference in the percentage of high school graduates can explain about 37.7% of the gap in education between blacks and whites.

These results are consistent with studies using social experiments in residential mobility. Public school choice lottery results (Hastings, Kane and Staiger, 2006) also show that the gains of students' test scores after attending their chosen school depends on the weight that their parents' preference for academic quality carries in their choices.

All of these suggest that there could be two selection processes and that neighborhood effects vary across the types.

The individual and household level data in this chapter comes from the National Longitudinal Surveys of Youth 1979 (NLSY79). The majority of county level information is from NLSY79; I supplement that data with county level information derived from either the US Census or from Centers for Disease Control and Prevention (CDC). The key neighborhood characteristic variable used in the chapter is race-specific county level education.

The remainder of the chapter is organized as follows: 2.2 is the data description. 2.3 first estimates the Heckman sample selection model, then introduces the type-consistent model, and proposes a solution to the endogeneity of the neighborhood variable. 2.4 estimates the model and presents the main empirical results of this chapter. And 2.5 concludes chapter with a discussion of the future research..

2.2 Data and descriptive statistics

The data for the analysis are primarily from National Longitudinal Surveys of Youth 1979 (NLSY79), supplemented by the 5% census sample in 1980, CDC's Compressed Mortality Files, the Geographic Information System (GIS) based census data, and residential segregation indexes from the census. The NLSY79 includes a representative sample of 12,686 individuals. These individuals were 14-22 years old when they were first interviewed in 1979, which was also the time for them to choose whether or not they would need to acquire post compulsory education. These individuals were interviewed annually through 1994 and on a biennial basis until 2000. NLSY79 includes individual, family, and some county (neighborhood) characteristics data.

A. Key variables

The key individual level variables in this chapter include the highest grade completed for an individual after compulsory schooling years and his/her moving status between 1979 and 1982. The individual schooling years after compulsory schooling are constructed by using highest grade completed minus compulsory schooling. I use school years beyond compulsory schooling since this represents educational attainment determined by the individual. I assume that in 1979 individuals make two decisions: years of schooling after compulsory education, and whether they want to change their neighborhoods or not. To exclude the residential mobility resulting from entering college, I restrict the sample to individuals between the ages of 14 and 17. Individuals who drop out of school before they finish their compulsory education are also excluded from our sample. Keane and Wolpin (1997) use the schooling years after age 16 as a proxy for schooling beyond compulsory years when they study the schooling and career choices for young men.

Compulsory schooling years are derived from compulsory attendance laws and child labor laws. Compulsory attendance laws specify a minimum and maximum age between which attendance is required and the minimum period of attendance. The child labor law regulates the employment of minors and the minimum of age to work. Those laws differ across states, and child labor laws and compulsory attendance laws have different requirements for leaving school. Following Lleras-Muney (2005), I construct two measurements for schooling years: work permit age minus entrance age, and leaving age minus entrance age. I choose the minimum requirement of these as our variable for compulsory schooling years. The years of compulsory schooling vary from six years to ten years. In table A1.1-1, the average additional schooling years are 5.315 for whites and 4.93 for blacks.

In the 1979, 1980, 1982 and 2000 interviews, information from individuals' five most recent moves and residences was collected in NLSY79 at the county level. An individual is defined as a mover if his county of residence in 1979 is different from his county of residence in 1982; otherwise, he is defined as a non-mover. I do not have information on people who moved within a county. People who moved away first and then moved back to the same county are categorized as non-movers. In our sample, 22.21% of whites, 16.90% of blacks, and 14.77% of other races are movers.

Defining the neighborhood of an individual is very difficult. This chapter uses the county in which the individual resides and the individual's racial group as his neighborhood, because of data availability, and the substantial segregation between blacks and whites in metropolitan areas and counties. Since the focus of this chapter is the neighborhood effect on the highest grade completed, the ideal neighborhood variable would be average schooling years at the neighborhood level. However, data on the percentage of the population who finish high school and college are available only by county. I expect that those percentage variables are highly related to average schooling years. Therefore, the key neighborhood variable in this chapter is the percentage of population finishing high school.

However, only the overall percentage of high school graduates at the county level is available in the NLSY. To construct race-specific high school graduates at the county level, I use the1980 IPUMS (Integrated Public Used Microdata Series) 5% census samples to construct two data series: the metropolitan level percentage of high school graduates (S_{MA}) and the metropolitan level race-specific percentage of high school graduates ($S_{MA\ /Black}$ and $S_{MA\ /White}$). Then I use the ratio between the metropolitan level race-specific percentage of high school graduates and the metropolitan level percentage of high school graduates ($S_{MA\ /Black}/S_{MA}$ and $S_{MA\ /White}/S_{MA}$) to construct the county level race-specific percentage of high school graduates by using county level percentage of high school graduates multiplied by the above ratio. This strategy is based on the assumption that the ratio for the percentage of high school graduates is the same for the county and the metropolitan area.

Appendix B Figure 1 shows that there is a positive relationship between the individual education and the neighborhood variables, i.e., the county-level race specific percentages of high school graduates, for both movers and non-movers. The slope is larger for the mover group than for the non-mover group, suggesting that moving may induce a difference in the neighborhood effect. The simple regression results, listed below in figure 1, reveal the same relationship. The slope estimates are statistically significant at 0.0297 for the movers and 0.0210 for the non-movers

Table A1.1-2 lists the county-level overall percentages of high school graduates for both whites and blacks, as well as the race-specific percentages of high school graduates for both whites and blacks. Overall, blacks live in counties that have lower percentages of high school graduates. The difference is even larger if it is race-specific, indicating that black neighborhoods have an even lower percentage of high school graduates.

B. Other variables

Table A1.1-1 lists other individual characteristics variables, including race, gender, age, residence (at the county level), religion (1 if religion frequency is more than once per week), and feelings about the safety of school (1 if the individual feels safe). Variables about family background include mother and father's education, family size, the oldest sibling's highest grade completed, family income, whether or not the individual lived with his parents until the age of 18, and family poverty status.

In addition to the individual and family background, table A1.1-2 presents a set of county characteristics as county level control variables. These variables are referred in Manski (1993) as contextual variables. These variables include the race-specific mortality risks, obtained from the CDC. In particular, the external mortality rate for ages 15 to 34 is used, under the assumptions that a huge disparity of mortality risks between black and white teenagers is due to external reasons such as homicide and that the peer group for teenagers is the young group with similar ages. During this time period, blacks' external mortality rate is about 162 per 100,000 people, while whites' is only 89.2 per 100,000 people.

Other county characteristics variables include the percentage of blacks in the population, the percentage of people living in urban areas, arrest rates, employment rate in the education institution, and the overall unemployment rate. Table A1.1-2 lists the means and standard deviations for the key variables for the whole population and by race. Neighborhood variables also differ significantly. Blacks also have a much higher crime rate; theirs is 5,833 per 100,000 people verses the white crime rate of 4,949 per 100,000.

Table A1.2 presents differences in county-level variables between movers and non-movers in 1979 (before moving) and 1982 (after moving) by race. The number in the table is percentage difference between movers and non-movers. For example, for the mortality rate, the number in column (1) is -1.1298, indicating that black movers' county mortality rate is 1.1298% lower than the black non-movers county mortality rate. Since the number in column (2) (after they move) for mortality risk is -10.8796%, blacks move from counties with slightly less mortality risks to counties with significant less mortality risks. In summary, table A1.2 shows: (a) Before moving, compared with white non-movers, white movers' neighborhoods have a lower percentages of blacks, lower unemployment, lower education levels, lower incomes, and less crime rates; (b) Compared with black non-movers, before their moving, black movers used to live in neighborhoods with fewer blacks, lower educational levels, and less crime known to the police; (c) For both whites and blacks, it is not difficult to find that movers usually move to neighborhoods with lower external mortality rates, higher percentages of urban population, and higher education levels; (d) The comparison of movers' and non-movers' neighborhoods before moving and after moving reflects a process of racial integration, at least at county level. Black movers usually move to the neighborhood with a lower percentage of blacks, and whites' new neighborhoods have a higher percentage of blacks. Blacks' new neighborhoods have higher income and lower unemployment, but this is not the case for whites. The new neighborhoods for both races have more crimes known to the police.[1]

C. Data imputation and sample size

A substantial number of county and individual variables are involved in this chapter. As a result, the intersection of non-missing values of all these variables is a rather small sample size. For example, from NLSY79, there are 4,963 individuals who are between the ages of 14 and 17, with positive schooling years and a non-missing record for the residence in 1979-1982. However, the intersection of all non-missing variables reduces the sample size to 918. Therefore, it is necessary to impute missing variables to preserve the sample size to some extent.

I use the mean imputation method. This method uses the mean of non-missing values to impute the missing values. I have three levels of variables: individual level, county level, and state and metropolitan level. The following imputation rule is adopted: (1) For the individual level data, I use race-specific national mean data to replace the missing data. (2) For the county level data, I use race-specific state mean data. (3) For the state level or metropolitan level data, I use race-specific national mean data. In addition to imputation, I also construct a missing dummy variable for each variable with missing observations. The variable takes on the value of one if an observation is missing, and zero otherwise. Among the total 4,963 observations, 3,686 observations remain after imputation. The rest of the missing values are from the variables I do not impute, most

[1] The fact that 1982 crime rate and arrest rate are much higher than 1979 is consistent with the 1980's increase in the crime rate and arrest rate.

of which are dummy variables. Table C1 presents the data statistics before imputation and after imputation.

2.3 The model

A. Choice of schooling years

Suppose that individual i is a member of group g. He (with his parents) chooses the schooling years S_i after his compulsory schooling. Let S_i be linearly related to the following control variables: (i) individual level characteristics X_i; (ii) group level variables that are predetermined at the time that choices are made, Z_g; (iii) an individual's perception of the average choice of others, S_g ; and (iv) unobservable individual and group attributes v_i and η_g, which make up the error term. According to Manski (1993), (ii) are the contextual effects, (iii) are the endogenous effects, and (iv) are the correlated effects. Therefore, the basic equation of schools years can be written as

$$S_i = \beta_1 X_i + \beta_2 Z_g + \beta_3 S_g + v_i + \eta_g \,. \tag{1.1}$$

Non-random sorting processes may create correlation between S_g and the error term v_i. For example, those people with positive v_i are likely to move to areas or neighborhoods with high quality schools. This may create a correlation between the observed group characteristics such as percentage of high school graduates enrolling in colleges.

B. Residential mobility choice

The residential mobility choice model in this chapter is given by

$$move_i = 1(\delta_o Y_{og} + \delta_d Y_{dg} + \delta_1 X_i + e_i > 0) \equiv 1\left(z_{ig}\delta + e_i\right) \quad (1.2)$$

where $move_i$ is 1 if individual *i*'s location at 1979 is different from his location at 1982, and is equal to 0 otherwise. The moving decision depends on the individual characteristics X_i, the contextual variables in the current location Y_{og} , and the destination Y_{dg}. In this chapter, the current location, *o*, is the county where the individual resided in 1979 and the destination, *d*, is the county where he resided in 1982.

Moving can be driven by a multitude of reasons: convenience, cost saving, or random shocks to the family or the job. This chapter investigates the relationship between the moving and education investment decisions; therefore, children's welfare is assumed to be one of the factors the parents will consider when they choose their location.

Since I can only observe whether an individual moves and information on the origin and destination, and no information relating to the steps involved in search is observed, the estimation of the above equation for alternative location is impossible. However, ignoring those alternatives may cause the estimation to be inconsistent. Following Ioannides and Zanella (2007), I choose the area-level characteristics of spatially adjacent counties as instrumental variables. There are two reasons to use these variables as instruments. First, if different areas are characterized by different distributions, Weitzman (1979) shows that the optimal search strategy is the nested strategy in which people search among areas first, and then search within the area. This type of search strategy suggests that average characteristics of adjacent counties would be correlated with the own-county characteristics. Therefore, using the average characteristics of adjacent counties as instrumental variables would be appropriate. Moreover, even if a person's search strategy is not the nested strategy of Weitzman (1979), it is still appropriate to use the adjacent counties as instruments as long as these instruments are well-correlated with the own-county characteristics, and are uncorrelated with the error term of the main equation. It is worth noting here that the instrumental regression does not require that I correctly specify the relationship between the instrumental variables and the endogenous variable; it only requires a correlation between these variables.

The construction of the instrumental variables in this chapter is slightly different from Ioannides and Zanella (2007). They choose the averages of the characteristics within the entire group of spatially adjacent census tracts, including the census tract where households reside as their instrumental variables. Including the location where the household resides could introduce some correlation between the instrument variables and the error term. Therefore, in calculating the spatially adjacent county means, I exclude the individual's county of residence.

C. Joint decisions of mobility and schooling

Here I allow the schooling decision and residential mobility to be jointly determined. The residential mobility decision is described in equation (1.2). The schooling-choice model with mobility is assumed to be the following:

$$S_i = \beta_1 X_i + \beta_2 Z_{og} + \beta_3 S_{og} + \beta_4 S_{og} * move_{ig} + v_i + \eta_g \,. \qquad (1.3)$$

In equation (1.3), I allow the neighborhood effect of current location, *o*, to affect the schooling choice. Figure 1 suggests that movers and non-movers may differ in their responses to the neighborhood variable. This difference may arise from the fact that those who move are more sensitive to the neighborhood than those who do not move. The parameter β_4 captures the differential effect of S_{og}. I use the 1979 Z_{og} and S_{og} values in this equation.

The endogenous moving process means that unobserved factors driving individuals' moving also have impacts on individual' schooling choices. If the moving decision is modeled as in (1.2), then the error term e_i in equation (1.2) is correlated with the error term v_i in equation (1.3). Assuming that Corr(v_i, e_i) = ξ, and taking the conditional expectation for movers and non-movers, I have

$$E\left(S_{ig} \mid move_{ig} = 1\right) = \beta_1 X_i + \beta_2 Z_{og} + \left(\beta_3 + \beta_4\right) S_{og} + \zeta \left(\frac{\phi\left(Z\delta\right)}{\Phi\left(Z\delta\right)} \right) \tag{1.4.1}$$

$$E\left(S_{ig} \mid move_{ig} = 0\right) = \beta_1 X_i + \beta_2 Z_{og} + \beta_3 S_{og} + \zeta \left(\frac{-\phi\left(Z\delta\right)}{1-\Phi\left(Z\delta\right)} \right) \tag{1.4.2}$$

Equations (4.1) and (4.2) are the Heckman sample selection model. Equation (4.1) is for the movers and equation (4.2) is for the non-movers. Testing the statistical significance of the coefficient ξ in equation (4.1) using the sub-sample of movers is a test for the endogeneity of the moving decision. Similarly, one can use the sub-sample of non-movers to test the statistical significance of the coefficient estimate of ξ in equation (4.2).

A less well-known implication of (4.1) and (4.2), however, is that the coefficient for the Heckman correction term $\varphi(z\delta)/\Phi(z\delta)$ for movers is the same as the coefficient for the term $-\varphi(z\delta)/(1-\Phi(z\delta))$ for non-movers. Therefore, testing the equality these two coefficients can serve as a specification test of the model.[2]

I use the two-step Heckman method to estimate the model. In the first step, I estimate the residential mobility choice model described in equation (1.2). Before I estimate the moving probit model, I apply the Rivers and Vuong (1988) procedure to test the endogeneity of county variables for the original place and the destination. Table A1.6 presents the results. As discussed earlier, consistently estimating this model requires a set of instrumental variables. I use adjacent counties information as the IVs. Table A1.5 presents the estimation results from the first step with the set of instrumental variables. The estimated coefficients, $\hat{\delta}$, are used to construct the inverse mills ratio term $\phi(z\hat{\delta})/\Phi(z\hat{\delta})$ for the mover sub-sample and the term $-\phi(z\hat{\delta})/(1-\Phi(z\hat{\delta}))$ for the non-mover subsample. Table A1.3 lists the estimation results for the Heckman sample selection models for both movers and non-movers. For comparison purposes, the OLS results are also presented in table A1.3 for both movers and non-movers.

It is clear from table A1.3 that the coefficients between movers and non-movers are different, while the coefficients between the OLS and the Heckman sample selection model are only marginally different. It is important to point out that the coefficients for the neighborhood variable—the county level race-specific percentages of high school

---

[2] In fact, the coefficient of X_i should also be the same for both equations. Testing the equality of the coefficient β can also serve as a specification test.

graduates—are insignificant for both the mover subsample and the non-mover subsample in the Heckman model. Although the coefficient is marginally significant (at the 10% significant level) for the non-mover subsample in OLS, the Heckman sample selection test rejects the OLS model for this subsample. Therefore, despite an unconditional positive relationship between the percentage of high school graduates and schooling years illustrated in figure 1, this neighborhood variable does not have any effect on schooling years after controlling for other variables in the Heckman sample selection model.

More important, the Heckman sample selection test exhibits inconsistent results from two subsamples. As discussed earlier, the coefficients from the inverse mills ratios should be the same for both subsamples. For the non-mover group, the coefficient estimate for the inverse of mills ratio is a statistically significant value of 2.046 (0.773), suggesting that the neighborhood choice is endogenous. However, if I use the mover group, the coefficient estimate for inverse mills ratio is statistically insignificant at the value of 0.604 (1.285), suggesting that the moving decision is exogenous. This difference suggests the possibility that the population may be heterogeneous in its motivations for moving. In the next subsection, I consider an extension of the current model in which the moving decisions may either be endogenous or exogenous to the schooling choice.

D. A model of heterogeneous motivation in moving

People move for multiple reasons. For example, they may move for better economic prospects. Bowles (1970) finds that economic incentives affect geographic mobility of workers. By estimating a dynamic model of search, Kennan and Walker (2006) conclude that differences in expected returns are important in driving migrations across the United States. Borjas, Bronars, and Trejo (1992) also make a similar conclusion. Other papers suggest that mobility may be driven by the individual taste and characteristics. For example, college-educated couples usually locate in larger cities to pursue more job opportunities, better quality of life, and a business environment (Chen and Rosenthal, 2006; Costa and Kahn, 2000). There are also Schelling-type motives; i.e., white American families tend to move out of areas where the share of minorities is above a critical point (Card, Mas and Rothstein, 2007).

Therefore, it is possible that residential mobility can be either endogenous or exogenous to the schooling choice. I assume there are two types of people: one is the endogenous type and one is the exogenous type. Because these two types of people may have different preferences, it is possible that their neighborhood effects may affect their schooling choices differently. Consider an extension of equation (1.3):

$$S_i = \beta_{11} X_i + \beta_{12} Z_o + \beta_{13} S_o + \beta_{14} S_o * move_i + v_{1i} + \eta_o \qquad (1.5.1)$$

$$S_i = \beta_{01} X_i + \beta_{02} Z_o + \beta_{03} S_o + \beta_{04} S_o * move_i + v_{0i} + \eta_o \qquad (1.5.2)$$

For equations (1.5.1) and (1.5.2), the coefficients for $X_{i,}$ Z_o, S_o are allowed to be different to capture the possible behavioral differences for the two types of individuals. Without loss of generality, I let the moving decision for type 1 in equation (1.5.1) be endogenous, and let the moving decisions for type 0 in equation (1.5.2) be exogenous. Given that the moving decision is modeled in equation (1.2), I have

$$\text{Corr}(v_{1i}, e_i) = \xi, \text{ and } \text{Corr}(v_{0i}, e_i) = 0. \tag{1.6}$$

Further, suppose that an unobserved variable T_i governs individual i's type: When T_i = 0, individual i is type 0 (exogenous type); when T_i = 1, individual i is type 1 (endogenous type). Therefore, the observed education S_i can be written as

$$S_i = (1-T_i)\, S_{0i} + T_i\, S_{1i.} \tag{1.7}$$

The type variable T_i is assumed to be determined by a vector of family and individual characteristics variables, denoted as w_i.

$$T_i = 1(w_i\pi + \tau_i > 0) \;\text{ where }\; \tau_i \sim N(0,1) \tag{1.8}$$

I assume that covariance of between (v_{0i}, τ_i) and (v_{1i}, τ_i) is zero in this analysis. After plugging S_0 and S_1 into the S_i in equation (1.3), I have the following model on schooling choice:

$$\begin{aligned} S_i = {} & T(\beta_{11} - \beta_{01})X_i + T(\beta_{12} - \beta_{02})Z_o + T(\beta_{13} - \beta_{03})S_o + T(\beta_{14} - \beta_{04})S_o * move_i \\ & + \beta_{01}X_i + \beta_{02}Z_o + \beta_{03}S_{oi} + \beta_{04}S_o * move_i \\ & + \lambda_i + \eta_o + \varepsilon_{io} \end{aligned}$$

where $\lambda_i = Tv_{1i} + (1-T)\, v_{0i.}$

Because only type 1 individuals are assumed to engage in endogenous moving while type 0 individuals are assumed to move for exogenous factors, the expectations conditioning on types and moves are given by

$$E(v_i \mid move_i, type = 1) = \zeta\left[(1 - move_i)\frac{-\phi(z\delta)}{1 - \Phi(z\delta)} + move_i\frac{\phi(z\delta)}{\Phi(z\delta)}\right]$$

$$E(v_i \mid move_i, type = 0) = 0$$

Since types are unobserved, the expectations conditioning on moves only are given by

$$E(S_i \mid move=1) = E(S_{1i} \mid move=1, T=1)\Pr(T=1 \mid move=1) + E(S_{0i} \mid move=1, T=0)\Pr(T=0 \mid move=1)$$

$$E(S_i \mid move=0) = E(S_{1i} \mid move=0, T=1)\Pr(T=1 \mid move=0) + E(S_{0i} \mid move=0, T=0)\Pr(T=0 \mid move=0)$$

Therefore,

$$\begin{aligned} E(S_i \mid move=1) = &\left[\beta_{11}X_i + (\beta_{12})Z_o + (\beta_{13}+\beta_{14})S_o\right]\Phi(\omega_i\pi) \\ &+\left[\beta_{01}X_i + (\beta_{02})Z_o + (\beta_{03}+\beta_{04})S_o\right](1-\Phi(\omega_i\pi)) \\ &+\zeta\left[\frac{\phi(z\delta)}{\Phi(z\delta)}\right]\Phi(\omega_i\pi) + \eta_o \end{aligned} \tag{1.9.1}$$

$$\begin{aligned} E(S_i \mid move=0) = &\left(\beta_{11}X_i + \beta_{12}Z_o + \beta_{13}S_o\right)\Phi(\omega_i\pi) \\ &+\left(\beta_{01}X_i + \beta_{02}Z_o + \beta_{03}S_o\right)(1-\Phi(\omega_i\pi)) \\ &+\zeta\left[\frac{-\phi(z\delta)}{1-\Phi(z\delta)}\right]\Phi(\omega_i\pi) + \eta_o \end{aligned} \tag{1.9.2}$$

Equations (1.9.1) and (1.9.2) together describe our type-consistent model. Comparing equation (1.9.1) with the Heckman sample selection model of (1.4.1), and (1.9.2) with equation (1.4.2), I find that our model is reduced to the Heckman model in two circumstances. First, if there is only one type of individuals, our model becomes the Heckman sample selection model. In the case that all individuals belong to the exogenous type, $\Pr(T_i=0)=1$, i.e., $\Phi(w_i\pi) = 0$, (1.9.1) and (1.9.2) become (1.4.1) and (1.4.2), respectively, without the Heckman sample correction term. In the case that all individuals belong to the endogenous type, $\Pr(T_i=1)=1$, i.e., $\Phi(w_i\pi) = 1$, (1.9.1) and (1.9.2) become (1.4.1) and (1.4.2), respectively, with the Heckman sample correction term. Second, when the coefficients of two types are the same; i.e., the behavior responses of the two types of individuals are the same, equation (1.9.1) and equation

(1.4.1) only differ by the probability term $\Phi(w_i\pi)$. This is intuitive since the model assumes that only part of the people belong to the endogenous type. The probability term $\Phi(w_i\pi)$ precisely gives the proportion of people who are endogenous. Estimates using (1.4.1) would overestimate the coefficient for the Heckman sample correction term.

The type-consistent model suggested in this chapter belongs to the family of the mixture density models. The mixture density models have been widely used in the literature to model unobserved types. For example, Feinstein (1990) proposes and estimates a mixture density that considers the unobserved violations and the observed detections of violations of laws and regulations. Keane and Wolpin (1997) model five different unobserved types in ability endowment using a mixture density model. Knittel and Stango (2003) estimate a mixture density model using state-mandated price ceilings as focal points for unobserved tacit collusions of credit card companies. When modeling the default probability of consumer credit cards, Gan and Mosquera (2008) suggest that different types of consumers may come from the heterogeneity in consumers' time discount rates, and/or the heterogeneity in their risk aversion. A model with unobserved types would lead to better out-of-sample predictions in default probability. Identification of this type of model, however, depends on the assumptions of the distribution. Here I assume that error distributions are normal since there are no compelling reasons to assume any other distributions.

Estimation of (1.9.1) and (1.9.2) requires two steps. First, I estimate the residential mobility model to obtain the inverse mills ratio term. The results of this estimation are listed in Table A1.5. In the second step, I estimate both (1.9.1) and (1.9.2) simultaneously. The parameters to be estimated include parameters from the schooling choice models, β, and the parameters from the type determination model, π. Nonlinear least squares is applied to equation (1.9.1) for movers and (1.9.2) for non-movers simultaneously.

Since the type-consistent model is highly non-linear, it useful to understand its small sample properties. In the Appendix D, I conduct a simulation study with sample sizes of 100, 500, 1,000 and 3,000. It is shown that when the sample size is below 1,000, large biases may occur and estimates are no longer reliable. When the sample size is 1,000, some biases may occur and estimates must be taken with caution. The estimates are very close to the true values when the sample size is 3,000.

E. Endogeneity of the neighborhood variable and its IV

In the education demand equation, the perceived neighborhood education variable S_o is possibly related to the error term η_o; thus, I must find appropriate instrument variables. The chapter uses instruments similar to those used in Ioannides and Zabel (2008): the county means of the inverse mills ratios. The county means of the inverse mills ratios are constructed from county variables and derived from the econometric assumption. Therefore, the county inverse mills ratio will not be correlated with unobserved county characteristics, and it is a valid instrumental variable (Brock and Durlauf, 2001; Manski, 1993; Moffitt, 2001).

Another source of instruments is the racial residential segregation indexes. These indexes are calculated at the metropolitan level. Residential segregation indexes usually measure the degree of racial segregation in terms of residence in five dimensions:

evenness, exposure, concentration, centralization, and clustering. The assumption here is that a black's neighborhood would be the black population in the county where he resides, and a white's neighborhood would be the white population in the county where he resides. However, the degrees of residential segregation may be different not only across the counties, but also within the counties. Therefore, our neighborhood variable can only be a proxy of the true neighborhood variable. In other words, our neighborhood variable is a measure of the true neighborhood variable with measurement errors, causing an endogeneity problem. To mitigate this problem, I use the segregation indexes as instrumental variables, which should mitigate these biases. Obviously the segregation indexes would be correlated with our neighborhood variable, but there is no reason to believe that these variables would have an effect on schooling in addition to their effect on the neighborhood variable. This chapter uses seven different measures of the segregation indexes as instrumental variables for our key neighborhood variable, the race-specific percentage of high school graduates. These seven measures capture all five dimensions of segregation. Table C2-1 describes these seven indexes in detail, and table C2-2 lists their summary statistics. All these indexes are directly obtained from the Census Bureau.

I use a weak identification test to test the above sets of instruments, and the Cragg-Donald test statistic is 95.434, much larger than the Stock-Yogo weak IV test critical values for 5% relative bias (20.53) and 10% size (36.19). Therefore, I reject the null hypotheses of weak identification. Table A1.4 lists the first stage results for the percentage of high school graduates.

2.4 Estimation results

The estimation involves several stages. First, I estimate a probit model for the moving decision by using the adjacent county means as instrumental variables. Second, using the county mean inverse mills ratio derived from the moving estimation and segregation indexes as IVs for the endogenous neighborhood variable, I get the predicted race-specific percentage of high school graduates at county level. Third, the predicted county education level is used to estimate the type-consistent model. In particular, I simultaneously estimate the type determination equation and the neighborhood effects on the education equations for the two types.

A. Residential mobility choice estimation

The residential mobility choice is modeled in equation (1.2). The contextual variables for both their current location Y_{og} and their destination Y_{dg} used in this chapter include percentage of population living in urban areas, percentage of households with a female as the household head, the crime level, percentage of high school graduates, percentage of college graduates, the unemployment rate, the total employment, and per capita income. All variables are at the county level. The individual and family background variables include the father and mother’s highest grades achieved, family size, age and age squared, a dummy for blacks, a dummy for male, and a dummy for living in an urban area. If a variable has missing values, I construct a missing dummy for those observations that have missing values.

As discussed earlier, the contextual variables are potentially endogenous. The strategy is to employ the mean value of adjacent counties (excluding the county where the individual resides[3]) as instrumental variables.

To test the endogeneity of the contextual variables, the Rivers and Vuong (1988) two-stage procedure is implemented. First, I run each neighborhood characteristic variable on its instruments (the mean of adjacent counties) and get residuals; second, I run the probit model with individual variables, all the neighborhood characteristic variables, and all the residuals from the first stage. If the coefficients before the residuals are significant, then corresponding variables are endogenous; otherwise, they are exogenous. The second stage results are presented in table A1.6**.** From this table, only the civil unemployment rates at the origin and the destination and the per capita income at the origin are endogenous. Therefore, I only instrument for these three endogenous variables in our estimation. The estimation results are presented in Table A1.5.

Before I discuss the estimation results, I follow Ioannides and Zanella (2007) to label a variable as an "attractor" if its coefficient is negative for the current county, and positive for the destination county. Similar, a variable is labeled as a "repeller" if its coefficient is positive for the current county, and negative for the destination county. I expect that for the same variable the coefficients for the origin and the destination will have opposite signs and same absolute value.

$$H_1:\ \text{sgn}(\delta_{o,j}) = -\ \text{sgn}(\delta_{d,j}) \qquad \text{for all } j.$$

$$H_2:\ \delta_{o,j} = -\ \delta_{d,j} \qquad \text{for all } j.$$

[3] The adjacent counties for each county are found from GIS programming and *I* thank Ms Yige Gao's help and guide in our GIS programming.

In table A1.5, I find that percentage of urban population, crime rate, percentage of high school graduates, percentage of college graduates, unemployment rate, and per capita income have significant impacts on people's moving decision. Among the above variables, the percentage of urban population, the percentage of high school graduates, and the percentage of college graduates are all attractors in individuals' moving decision. People would like to stay in places with a high percentage of urban population and a high education level. On the other hand, the unemployment rate acts as a repeller. People would like to move away from places with high unemployment rates and move to places with high employment rates. The estimation results show that low crime rates and high per capita income will increase the probability of moving out. The estimation result for the crime rate is consistent with the fact that movers move to the residence with high crime rate (see table A1.2) because of the in the huge increase in the crime rate in 1980's. In terms of per capita income, black movers move to the location with higher per capita income, while white movers do not. The estimation result for per capita income reflects that on the average movers move to the location with low per capita income because I have more whites than blacks in the sample.

In addition, table A1.5 also shows that the inclination of moving also differs across families. It is related to parents' education background and family size. In most cases, moving is easier for the smaller families and for highly educated parents. The estimation also shows that moving usually happens more frequently before individuals are 16.6 years older and becomes less when individuals are older than 16.6 years. Finally, there is no statistically significant difference between blacks and other races in their moving probabilities.

B. Results from the type-consistent model

In subsection 2C, I discuss the Heckman sample selection model for movers and non-movers and present the estimation results in table A1.3. I show that the Heckman sample selection model produces inconsistent results for the mover group and the non-mover group. In this section, I estimate our type-consistent model. To be more general, I allow both types to be potentially endogenous. By doing this, I can test if one type is exogenous while another type is endogenous. In particular, I rewrite (1.6):

$$\mathrm{Corr}(v_{1i}, e_i) = \xi_1, \text{ and } \mathrm{Corr}(v_{0i}, e_i) = \xi_0 . \tag{1.10}$$

A specification test of our type-consistent model is that $\xi_1 \neq 0$ and $\xi_0 = 0$. Alternatively, a specification test for the Heckman sample selection is that $\xi_1 = \xi_0 = \xi$. In addition, it is entirely possible that $\xi_1 \neq \xi_0$, and neither one of them equals to zero. This suggests the existence of two endogenous types. These two types may exhibit different behavior in their education equations. With this setup, equations (9.1) and (9.2) are modified in the following way:

$$\begin{aligned} E(S_i \mid move = 1) = &\left[\beta_{11} X_i + (\beta_{12}) Z_o + (\beta_{13} + \beta_{14}) \hat{S}_o + \zeta_1 \frac{\phi(z\delta)}{\Phi(z\delta)} \right] \Phi(\omega_i \pi) \\ &+ \left[\beta_{01} X_i + (\beta_{02}) Z_o + (\beta_{03} + \beta_{04}) \hat{S}_o + \zeta_0 \frac{\phi(z\delta)}{\Phi(z\delta)} \right] (1 - \Phi(\omega_i \pi)) \end{aligned} \tag{1.11.1}$$

$$E(S_i \mid move = 0) = \left(\beta_{11}X_i + \beta_{12}Z_o + \beta_{13}\hat{S}_o + \zeta_1 \frac{-\phi(z\delta)}{1-\Phi(z\delta)}\right)\Phi(\omega_i\pi) + \left(\beta_{01}X_i + \beta_{02}Z_o + \beta_{03}\hat{S}_o + \zeta_0 \frac{-\phi(z\delta)}{1-\Phi(z\delta)}\right)\left(1-\Phi(\omega_i\pi)\right) \quad (1.11.2)$$

In both equations, $\hat{S}_o$ instead of S_o is used where $\hat{S}_o$ the predicted value from the instrumental variable regression. As discussed earlier, the instrumental variables are averages of the inverse mills ratios. Since the parameters δ and π are the same for both equations, it is necessary to simultaneously estimate both equations (1.11.1) and (1.11.2). The non-linear least squares method is applied to estimate this model.

Tables A1.7-1 and A1.7-2 present the estimation results. Table A1.7-1 reports the type estimation results, where the basic type determination model is described in equation (1.8). The coefficients reported from this table are similar to those from a binary probit-type of model. A positive coefficient of a variable indicates a positive marginal effect of that variable's contribution toward the endogenous type. Different from a regular binary probit model, here the types are unobserved.

The estimation results show that endogenous type people are more likely to be those who are from a larger family, who leave their parents before age 18, who are from a poor family, who do not feel safe about their school, and who are single. Males are more likely to belong to the endogenous type than females. Individuals whose eldest sibling does not have a high education level are more likely to be endogenous type. In addition, individuals whose parents are highly educated are more likely to be exogenous type. An interesting result here is that the race does not have a significant effect on determining the type of an individual.

Table A1.7-2 reports the estimation results for the schooling equations for both types. The first row in table A1.7-2 reports the covariance between the error e_i (from moving equation) and the error v_{is} (from the schooling equation for type s). The covariance ξ_1 is estimated to be -0.67 (0.38), with a p-value of 0.08, and the covariance ξ_0 is insignificant (point estimate is 0.68, and the standard error is 0.74). This result suggests the existence of both the exogenous type (type 0) and the endogenous type (type 1).

The table also shows that the neighborhood effects are different for different types of households. The key neighborhood variable in this chapter is the race-specific percentage of high school graduates at the county level. The only statistically significant effect is for the movers who belong to the endogenous type. The coefficient estimate for the interaction term between percentage high school graduates and moving dummy is 0.051 (0.014). Compared with the Heckman model in table A1.3, not only is the estimate in the type-consistent model statistically significantly while the estimates in Heckman models are not, but also the magnitude from the type-consistent is much larger than estimates from other models.

From table A1.1-2, I know that the mean difference in percentage of high graduates between white neighborhoods and black neighborhoods is 21.5%. Therefore, if an endogenous type person moves from an average black neighborhood to an average white neighborhood, his education level would increase by 0.051* 21.5% = 1.0965 years. For the exogenous type person, the proportion of high school graduates does not have any significant impacts on his educational attainment. The result suggests that the neighborhood effect is concentrated on the movers who belong to the endogenous type.

Furthermore, individual, family background variables, and other county variables also show different impacts for different types. For the endogenous type, religion has a positive impact on educational attainment. Religious people will choose higher education levels. For the endogenous type, the percentage of blacks and the mortality rate also have positive impacts on educational attainment.

For the exogenous type, mother's education level has a negative impact on children's education. The unemployment rate has a positive impact on educational attainment. One plausible explanation is that the high unemployment rate will give individuals more incentive to study hard, since higher education lowers their risk of becoming unemployed. Another possible reason is the high unemployment rate makes finding work difficult, so people stay in school longer than they would have when unemployment rate was low.

In conclusion, the above results imply that neighborhood effects have differential effects on educational attainment. If the location choice is endogenous, there is a substantial neighborhood effect for movers. However, no statistically significant neighborhood effect exists if the location choice is exogenous. Furthermore, the two types of people behave differently. Some characteristics may affect the education choice for one type of people but not for the others. Given the fact that table A1.3, based on the Heckman selection model, shows no statistically significant neighborhood effect, it is important to account for the heterogeneity. It is worth pointing out here that our conclusions are consistent with what social experiments have found.

Based on the estimation results from table A1.7, table A1.8 calculates the numbers of endogenous types and exogenous types for movers and non-movers. Among the 3,121 non-movers, 81.35% (or 2,539) belong to the endogenous type. For the 565 movers, the percentage of belonging to the endogenous type is slightly higher, at 86.37% (or 488). Overall, 82.12% of the sample belongs to the endogenous type.

Table A1.8 lists the average years of schooling by type and moving status. Overall, movers have more years of schooling than non-movers. More important, the endogenous type has 1.83 years or 49.56% more education beyond compulsory schooling than the exogenous type. For the movers, the difference is even more striking. The endogenous type has 91.7% more schooling years than the exogenous type. This large difference in education between the two types is consistent with the notion that the types identify two distinct groups of people.

Table A1.9 lists the observed and predicted education level by type and race. Panel A has the observed schooling years while panel B presents the predicted the schooling years. It is clear from the table that the type-consistent model predicts schooling years at the mean level quite well. The largest difference is about 5.3% for the blacks who are exogenous type. However, the predicted standard deviations are larger than the observed standard deviations for all types and all races. This occurs because of the nonlinearity of the model.

As discussed earlier, if I let endogenous black movers have whites' neighborhood value where the percentage of high school graduates is 21.5 percentage points higher, these black movers' education level would increase by 1.0965 years. Table A1.8 shows that movers account for 15.33% of population and that endogenous type accounts for 86.37% of all movers. Therefore, the average gains if blacks have the same neighborhood variable would be 1.0965*86.37%*15.33% = 0.1452 years, which is 37.7% of the overall difference at 0.385 years between blacks and whites in sample. Therefore, I conclude that the neighborhood effect accounts for 37.7% of the overall education difference between blacks and whites.

2.5 Conclusions

This chapter makes three main contributions to the neighborhood effect literature. First, it contributes to the econometric solution to the neighborhood effect by proposing a new model to identify the neighborhood effect. The key feature of the proposed model is its treatment of the unobserved types. The chapter presents two reasons to justify its assumption of unobserved types. The first reason is more theoretical. The chapter argues that people move for different reasons. One reason is to search for neighborhoods that may help education. In this case, the moving decision is endogenous. However, it is also typical for people to move for reasons that are independent to education choice. In this case, the moving decision is potentially exogenous. The second reason is empirical. The test based on Heckman sample selection model shows that coefficients for inverse mills ratio are different between the mover group and the non-mover group.

Second, this chapter provides strong evidences of the existence of the two unobserved types. The estimation is conducted by allowing possibilities of both the one-type model and the two-type model. The data supports the two-type model. In addition, people from different types exhibit different behaviors. The endogenous type has 91.7% more schooling years (beyond the compulsory schooling) than the exogenous type.

Third, introducing the two unobserved types of individuals helps us to identify the neighborhood effect. When I use the Heckman sample selection model, I do not find any statistically significant neighborhood effects. However, the type-consistent model shows that the neighborhood effect only shows up for the endogenous mover group. Therefore, it is not surprising that the effect is not seen for the population with one-type model. I find that neighborhood difference between blacks and whites is an important factor affecting their educational gap. In particular, I find that the county-level race specific percentage of high school graduates can explain 37.7% of education difference between blacks and whites.

One of the limitations of this chapter is the definition of neighborhood. This chapter uses the county as the neighborhood. It is almost certain that such a definition of the neighborhood is too large. If I were able to use a sufficiently narrower geographic definition (census tract or zip code, or perhaps some information about individuals' social circle), it is possible that magnitude of neighborhood effects would be larger. Here, still because of the limited information I have, I cannot make a formal comparison by using different definitions of neighborhood. However, it is likely that our results provide lower bounds of the true neighborhood effect.

CHAPTER III

MARRY FOR MONEY OR MARRY FOR ROMANCE: AN EMPIRICAL STUDY OF FEMALES' CHOICES OF EDUCATION AND PARTNERS

3.1 Introduction

One of the key demographic features in the US over the last several decades is the rapid rise in women's educational attainment. The educational gap between females and males has been declining or even reversed. More than half of American college students are women and about 60% of college graduates are females (Becker and Posner, 2008).

This chapter tries to shed some insight by studying the motivations of women for education. An obvious return for more schooling is a higher wage in the labor market. In addition, the literature has provided another potential return for education: a better educated person may be able to meet a better partner in the marriage market. While the benefits from the marriage market apply for both genders, it is likely more important for women since married men on the average have higher labor force participation rates and higher income than married women.

There is a wide literature base about the effect of education on the quality of partners and it finds positive assortive mating in terms of education (Boulier 1984; Mare 1991; Cancian et al.1993; Behrman et al. 1994; Juhn and Murphy 1997; Weiss and Willis 1997). Little work, however, investigates whether the mating and the education investment will be different if the education decisions are driven by different incentives.

This chapter focuses on empirically identifying the motivation of education and its effect on the prospect of partners. The chapter first suggests an identification strategy to distinguish these two types of motivations for education. In particular, if the woman's education decision is endogenous to the partner choice, then pursuing marriage market return is her motivation for education (i.e., the woman is the "marry-for-money" type). However, if the woman's education decision is exogenous to the partner choice, then her education decision is driven by the labor market return (i.e., the woman is the "marry-for-romance" type). Second, this chapter proposes and estimates an empirical model that incorporates these two types of motivations in the market. The central empirical challenge is how to identify the type, since the type is unobservable. The chapter suggests modeling the unobserved types by a mixture density to characterize the unobserved types. The estimation results suggest that both motivations exist in the population. A woman would have a probability of 0.388 to be the "marry-for-money" type, and the "marry-for-money" type would take 0.855 years less education than the "marry-for-romance" type. Further, at the same education level, the "marry-for-money" type woman would be much more likely to have a better educated partner than the "marry-for-romance" type. Finally, I find that a woman's attitude toward a female's role in a family is the key factor that identifies the two types of women. A woman with a more traditional view on the female's role in a family would be more likely to belong to the "marry-for-money" type.

This result suggests a potential reason for a recent rapid rise in women's educational attainment. The changes of women's attitude about the women's status in the family and in the society are documented in the literature. With more and more women becoming career-oriented, they are more attracted by the financial benefit from the higher education. Therefore marriage consideration becomes less and less important in their education decision. The more marry-for-romance type women the society has, the higher the educational attainment the women will have.

3.2 Literature review

This chapter sits at the juncture of three strands in economics literature. The first strand is on school's positive impacts on the marriage outcomes. Women's education can influence the marriage outcomes in the following aspects: the first is the husbands' income. Highly-educated wives can become their husbands' assistants in their jobs, or they can maintain their husbands' physical and mental health and contribute to a high quality lifestyle (Benham 1974). Empirically, it is found the wife's education level has positive impacts on husband's earnings. The magnitudes of the impacts differ across the countries and regions (the marginal effects of female education on husband's earnings range from 2%-5%).[4] In addition, people have found that the wife's education would help her to have a husband with a higher level of education. By examining how the educational attainment of identical twins correlates to the educational attainment of their spouses, Behrman, Rosenzweig, and Taubman (1994) and Behrman and Rosenzweig (2002) found that an individual who receives an extra year of schooling relative to his or her twin marries a spouse with 0.3 years of additional education on average. Schwartz

[4] Wong (1986) for Hong Kong; Grossbard-Shechtman and Neuman (1991) for Israel; Scully (1979) for Tehran, Iran; Benham (1974), Lam and Schoeni (1994) and Jepsen (2005) for Brazil and United States.

and Mare (2005) report trends in educational assortive marriage from 1940 to 2003 and they find that well-educated people, especially college graduates, are more likely to marry each other than those with less education. The marriage between persons at the extreme distribution of education declined. Therefore, schooling has become a more and more important channel to enhance the prospect of marriage. The third marriage outcome that education has an impact on is the probability of being single. The ratio of the probability of remaining single for skilled (college and above) relative to less skilled (less than college graduates) women has had a dramatic decline over time. Compared with the cohort born in 1890, the probability of being single for the generation who was born in 1950 and attended college around 1970 dropped from 31% to 7.9% if the women graduated from the college. And the probability of being single for non-college counterparts dropped from 7.8% to 5.5% (Fernández and Fogli 2002).

The second strand of literature related to this work investigates how the type of matching influences the investment efficiency, allocation efficiency and welfare. Booth and Coles (2005) assume that there are two matching paradigms: one where partners marry for money and the other where partners marry for romantic reasons orthogonal to productivity or debt. They found that different types of matching generate different investment incentives and therefore have a real impact on the economy. Marrying for money generates greater investment efficiency; romantic matching generates greater allocation efficiency. Furthermore, it is shown that when people marry for money, the equilibrium implies the perfect positive assortive matching. It implies that when women marry for money, their educational attainment will help them to find highly educated husbands. It is also shown that in romantic matching, where the matching is orthogonal to real economic variables, female education rate is lower than that of marry-for- money matching. When such different matching regimes arise, it suggests that if the education is provided by the state rather than by parents, or capital markets are perfect and young adults finance their own college education, children become more independent from their parents, then the society will end up with more romantic matching.

The third strand of the literature related to this chapter is the mixture density models that have been used in literature to model unobserved types. For example, Feinstein (1990) proposes and estimates a mixture density that considers the unobserved violations and the observed detections of violations of laws and regulations. Keane and Wolpin (1997) model five different unobserved types in the ability endowment using a mixture density model. Knittel and Stango (2003) estimate a mixture density model using state-mandated price ceilings as focal points for unobserved tacit collusions of credit card companies. When modeling the default probability of consumer credit cards, Gan and Mosquera (2008) suggest that different types of consumers may come from the heterogeneity in consumers' time discount rates, and/or in their risk aversion. A model with unobserved types would lead to better out-of-sample predictions in default probability. The identification of this type of model, however, depends on the assumptions of the distribution.

Compared with the above literature, this chapter has the following features: First, this chapter investigates the effect of wives' education on their husbands' education levels. A good husband can be reflected in many aspects: the personality, the wealth level, the education level, and talents. The education level is generally a good signal. A well-educated husband usually has a good job or career and therefore good financial prospects. Also the education is generally helpful for developing and shaping healthy and harmonious personality and talents. In a conclusion, the education level can be an index reflecting the husband's comprehensive ability. When considering whether the women's education decision is partly driven by their marriage incentive, it can be the decision to go to high school, or the decision to go to college. Here the decision of going to college is considered, because having some college education can make larger difference in mating. I use NLSY1979 data and divide the whole female sample into two groups: one group with their highest grades less than or equal to 12 years and the other group with the highest grades more than 12 years. I find that the relationship between the husband's education level and the wife's education level is not linear and the relationship is different for two groups. Figure 2 in Appendix B graphically presents this relationship. For females whose education level is less than 12 years, an additional year of education will make their husbands' education level increase by 0.476 year. For females whose education level is more than 12 years, an additional year of education will make their husbands' education level increase by 0.582 year. This nonlinear and positive relationship between the wife's education level and the husband's education level implies that college education can enhance the probability of finding a highly educated husband.

Second, instead of assuming that the marriage consideration is one of motivations behind all women's education decision, this chapter allows that some women have different considerations in which their education investment is orthogonal to the marriage consideration. This chapter empirically analyzes the women's education decision and the marriage prospect under more general assumptions. Third, by assuming people could have or could not have the marriage consideration, this chapter suggests the existence of two types of women: one is those with the marriage consideration and the other is those without it. However, as the type is not observable, this chapter applies the mixture density model to identify different types by assuming that the unobserved type follows the normal distribution.

The remainder of the chapter is organized as the following: section 3.3 develops the empirical models; Section3.4 presents the data set used in this chapter. Results are presented in section 3.5 and section 3.6 concludes the chapter.

3.3 Empirical models

To address the female's education investment decision and the marriage prospect, two behavioral processes are considered: attending college or not and the educational prospect of the mate. I first discuss each of the two processes separately and then discuss them in a joint framework.

A. Decision to attend college

The college decision is modeled as following:

$$College = 1(\theta_0 + \theta_1 Z + \lambda > 0) \tag{2.1}$$

where Z is the vector of observed factors that may affect the decision of going to college including individual level and county level characteristics. And λ is the unobserved random variable, which is assumed to have a standard normal distribution.

Empirically, College = 1 if the highest grade > 12, and 0 otherwise. *Z* may include mother's highest grade, father's highest grade, living in an urban area or not, poverty status, and county level unemployment rate. Parents' education levels are expected to help children to pursue a higher education. Females from an urban area and rich families usually end up with a higher education.

B. The partner's education level

In this chapter, good husbands are measured by their education level. When the husbands' education levels are much higher than their wives', it reflects not only that husbands' potential income or wealth will be much higher than their wives' but also the husbands' comprehensive abilities are relatively higher. Here the education level of the mate, although potentially continuous, is modeled as a discrete choice as well. In particular, if a woman's husband has a college degree, or he has six more years of education, the women's husband is considered to be a good husband, and the variable Marriage = 1. Otherwise, Marriage = 0. For those highly educated women (whose education level is also college or beyond college), as long as their husbands go to college, I also claim that they marry good husbands because I assume the lifetime income level will not have much difference for the people whose education level is college and beyond college.

$$Marriage = 1(\beta_0 + \beta_1 X + \beta_2 educ + \zeta > 0) \qquad (2.2)$$

where the variable *educ* is the females' education level. And X is the set of control variables that affect the women's decision to select a highly-educated husband or a poorly-educated husband. The unobserved random error ζ is assumed to have the standard normal distribution.

The right-hand side of the equation includes the wife's characteristics at the time of marriage: wife is from an urban area or not (1 is from the urban), education level of the wife's mother, poverty status of the wife family's and wife's age. "*educ*" is wife's education level at the time of marriage.

According to the previous discussion, the wife's education level is helpful for her to find a good husband. Additionally, if the wife's parents have high education, it is good for her to marry a good husband too. Also a wife from an urban area will have more chances to meet highly educated men.

C. Type

I assume there are two types of matching. If a female's decision about education investment is related to her marriage decision, then her matching type is "marry for money" or strategic matching, which means at least part of her education investment is driven by the marriage market return. Empirically, this assumption is equivalent to Cov(λ, ζ) $\neq$0, and the *educ* variable is endogenous in the marriage equation (2.2). However, if a female's decisions about her education and marriage are independent of each other, which means her education decision is not driven by her marriage market return, then her matching type is "marry for romance" or romantic matching. In the empirical model, the correlation between the unobserved heterogeneities in each equation is zero. And zero correlation means "marry for romance". Furthermore, in this model specification, I also allow that different type people can have different behaviors.

As discussed before, given different motivations behind the college education decision, female could be divided into two types. One type is the "marry for money" type. Their education investment decision is partly driven by their motive to marry a highly educated husband. Their pursuing high education is to increase their opportunities to match with men with higher education. Another type is the "marry for romance" type. Those females' education decisions to go to college have nothing to do with their marriage considerations. It is possible that different incentives will lead to different marriage prospects. Consider an extension of equation (2.1) and (2.2):

$$Marriage = 1(\beta_{t0} + \beta_{t1}X + \beta_{t2}educ + \zeta_t > 0) \tag{2.3}$$

$$College = 1(\theta_{t0} + \theta_{t1}Z + \lambda_t > 0) \tag{2.4}$$

where the type variable $t = 0,1$

Coefficients in the equation (2.3) and (2.4) are allowed to vary across the types to capture the possible behavior differences for the two types of women. Without losing generality, it is assumed to be the type of marrying for romance if t=0 and the type of marrying for money if t=1.

For the type 1 who marries for money, the unobserved heterogeneity components from the education and the marriage are correlated, while for the type 0 who marries for romance, the heterogeneity components are independent. Therefore, the heterogeneity components are assumed to have the following linear relationship for the type marrying for money:

$$\zeta_1 = \rho\lambda_1 + v \text{, where } v \sim N(0, 1-\rho^2) \text{ and } \rho \neq 0 \tag{2.5}$$

If a female marries for romance, then

$$\begin{pmatrix} \lambda \\ \zeta \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}\right)$$

Furthermore, suppose that an unobserved variable Ti governs the individual i's type: When Ti = 0, the individual i is the type 0 (marry-for-romance type); when Ti = 1, the individual i is the type 1 (marry-for-money type). The type variable Ti is assumed to be determined by a vector of family and individual characteristic variables, denoted as w_i.

$$T = 1(w_i\pi + \tau_i) \text{, where } \tau_i \sim N(0,1) \tag{2.6}$$

Because the type is not observed, this model belongs to the family of the mixture density models. The mixture density models have been widely used in the literature to model unobserved types. The identification is based on the assumption of distribution. Here τ_i is assumed to follow normal distribution.

D. Estimation

The estimation in this chapter will be based on the maximum likelihood function. According to the values of *College* and *Marriage*, four cases have to be considered: (1) *Marriage* =1 & *College* =0;(2) *Marriage* =0 & *College* =0; (3) *Marriage* =0 & *College* =1 and (4) *Marriage* =1 & *College* =1.

After deriving all of joint densities for the above cases, the likelihood function can be written as (details are presented in Appendix E):

$$
\begin{aligned}
l = & \sum_{Marriage=1,College=0} \log\left(\Pr\left(Marriage=1, College=0\right)\right) \\
& + \sum_{Marriage=0,College=0} \log\left(\Pr\left(Marriage=0, College=0\right)\right) \\
& + \sum_{Marriage=1,College=1} \log\left(\Pr\left(Marriage=1, College=1\right)\right) \\
& + \sum_{Marriage=0,College=1} \log\left(\Pr\left(Marriage=0, College=1\right)\right)
\end{aligned}
$$

In this model, the very important null hypothesis is the existence of two types of women i.e. $\rho \neq 0$. If the null hypothesis is not rejected, then it gives us a strong evidence of the existence of two types of women. If the null hypothesis is rejected, then it means that all the women belong to only one type: either marrying for money or marrying for romance.

3.4 Data

The data for the analysis are primarily from National Longitudinal Surveys of Youth 1979 (NLSY79). The NLSY79 includes representative sample of 12,686 individuals. These individuals were 14-22 years old when they are first interviewed in 1979. These individuals were interviewed annually through 1994 and were interviewed on a biennial basis until 2000. NLSY79 includes the individual, family, and county characteristics data. It is a good data set for this research because they gathered information in an event history format, in which dates are collected for the beginning and ending of important life events such as education, employment, and marriage. The NLSY79 contains three sub-samples: a cross-sectional sample of representative of the civilian U.S. youth population; a supplemental sample of over-sample civilian Hispanic, black, and economically disadvantaged non-black/non-Hispanic U.S. youth; and a sample of the population ages 17–21 who were enlisted in the military. Following the 1984 interview, the military subsample was no longer eligible for interviews. Therefore I exclude the military subsample from my analysis.

Table A2.1 contains the summary statistics for observations used in the analysis. In the model, only the first marriage is considered. Since the information about wife's marital status for each survey year is available, the year of the first marriage can be identified. Therefore, females' highest grades, the residence being urban or not, the poverty status and the age when they got married can be identified after knowing the year of their first marriage. According to Table A2.1, the average age for the first marriage is around 23 years old and the average highest grade when getting married is 12.7. Only around 10% of females are from the families under poverty and 77% of women were living in an urban area when they entered into their first marriage.

College is a dependent variable in the education decision. It is a dummy variable. If a women's highest grade is greater than 12, which means she at least has some college education experience, and then the value of *College* is 1, otherwise it is 0. Table A2.1 shows that less than 50% of females have the college education experience.

Marriage is also a dummy variable. It reflects the comparison of the husband's education level and the wife's education level. If the value of *Marriage* is 1, then it represents that the husband's education level is at least 6 years higher than the wife's education level or her husband has at least a college education. If none of the above conditions are satisfied, then the value of *Marriage* is 0, which means the husband's education is not high. In the sample, only around 37% women find highly educated husbands.

Table A2.2 presents the distribution of going to college and finding a good husband. Around 42% of husbands' education level is not much higher than their wives' or they do not have college education. About 29% of college-educated women marry college-educated men. 8% of women who do not have a college education marry a highly-educated husband. There are around 20% of women whose husband's education level is less than theirs.

In NLSY79, the respondents were asked a few questions related to the female's attitude towards women's status in the family and career. One of the questions is: "It is much better for everyone concerned if the man is the achiever outside the home and the women takes care of the home and family." And there are four categories of answers: 1 strongly disagrees, 2 disagree, 3 agree and 4 strongly agree. The variable -- traditional attitude to the women's status is 1 if the answer is agree or strongly agree, otherwise it is 0. Around 36% of women's attitude about the roles of husband and wife is still traditional. Other questions are related to whether a man should share the housework, whether women are happier in traditional roles, whether wife's employment leads to more juvenile delinquency, and whether a working woman feels more useful than the one who does not hold a job. The answers to those questions are strongly related to women's attitude about the roles of husband and wife. However, the answer to the first question can best represent the female's attitude about the marriage.

Other variables, such as the parents' highest grade, the unemployment rate in the county where the women reside, the region of their residence and whether they lived with their parents at age 14 are included in the analysis. Those variables are expected to have some impacts on the females' education decision and the marriage quality. Parents' education level is supposed to have positive impacts on their child's education and marriage prospect. On the average, parents finish the compulsory education and the highest grade is between 10 and 11. More than one-fourth of women are from north-central and more than one-third of women are from south and less than one-fourth of women are from north east and west.

Dummy variable "religious" is constructed as: if a woman attended the religious service at least once a week, then she is considered to be religious and the value for this variable is 1, otherwise it is 0. In the sample, around 36% of women are religious.

3.5 Results

First, equation (2.3) and (2.4) are estimated separately. Table A2.3 and A2.4 present the probit estimation results for equation (2.3) and (2.4). All of the estimations use the standardized value i.e all of variables are standardized by their own standard deviation. As we can see, parents' education has a positive impact on the female's education and the marginal effect is 0.0842 for mom's education and 0.128 for dad's education. The urban residence is also helpful for women to pursue college education and its marginal effect is 0.057. Family poverty status and county level unemployment rate have expected signs for women's college education but they are insignificant.

The estimation results for the marriage equation show that mom's education is helpful for women to find a good husband. If mom's education increases one standard deviation, their daughter's possibility of finding a good husband will increase by 0.0458. Urban residence has similar impacts on women's marriage prospect but with a little bit larger marginal effect at the value of 0.0623. As what expected, female's education level has comparatively large positive marginal effects on the possibility of finding a highly educated husband. The marginal effect is 0.232. It confirms the positive assortive mating between the husband's education and the wife's education. The result that women's age of getting married is positively related with the possibility of finding a well-educated husband is consistent with the fact that women's age of the first marriage increases.

Table A2.5 reports the estimation results if only marrying for money type exists. Column 3 is the MLE estimation result and column 4 is the probit results for comparison. First, if we ignore the type of marrying for romance and assume the whole group has only one type, i.e. the type of marrying for money, the estimated correlation between error terms from equation (2.3) and (2.4) is 0.33 and it is statistically significant. Comparing with the probit estimation for the education equation only, the estimation results from MLE are marginally similar, but the MLE results for the marriage equation is quite different from the probit results. For the education equation, mother's highest education and the urban residence also have positive impacts on their daughter's decision of going to college but with a slightly larger magnitude. The estimated coefficients for mom's education and the urban residence are 0.227 and 0.151; while the effects are only 0.211 and 0.143 if both equations are estimated independently. On the other hand, if women are the type of marrying for money, the effect of dad' highest education on their daughter's education is 0.3, a little bit smaller than 0.32 in the independent probit estimation.

For the marriage equation, the estimation results are much larger than the results from the probit estimation except for the women's age. When the correlation between these two equations is considered, the women's age of getting married is not significant any more. The estimated coefficient before mother's education is 0.257 for the MLE but only 0.125 for the probit model. Female's highest grade when getting married is much larger for the type of marrying for money with the value of 0.72 but only 0.63 for the probit model. It suggests that with marriage consideration, the female's education has larger marginal effects on the probability of finding a good husband, and therefore the positive assortive mating could be different for different types.

Table A2.6 reports the estimation results for the type model. First, the estimated value for λ is 0.32 and statistically significant. This value is quite close to the estimated value of 0.33 from the table A2.5. It gives a strong evidence of the existence of two types of women: the type of marrying for money and the type of marrying for romance. The positive sign of the variable suggests that any unobserved variable with positive impacts on the education will also have positive influence in finding a good husband. Second, it is worth noting that two types of women have different behaviors. The estimations for both the education equation and the marriage equation are quite different for two types. It means that different incentives will result in different marriage prospects and different education investment behaviors. Parents' education has a much larger effect on the daughter's education for the type of marrying for money than that for the type of marrying for romance. And mother's education plays a larger role for their daughter's marriage prospect if their daughters marry for money. These results imply that parents have a much stronger influence on the type of marrying for money. Comparing the results in table A2.5, it is not difficult to find out that the estimation results in table A2.5 are smaller than the estimation results for marrying for money, but larger than the results for marrying for romance. The Urban residence has positive impacts on women's education and also is helpful for women to find a good husband. The magnitudes of the estimations of the urban residence for both types and in both equations are not quite different. But the impact of the women's education level on the marriage prospect is quite different for two types. For the type of marrying for money, the women's education will largely enhance their probability of finding a highly educated husband. It will be also helpful for the women who marry for romance to find a college-educated husband, while the marginal effect is much smaller than that for the type of marrying for money. This result suggests that different types have different assortive matchings. The type of marrying for money has much larger positive assortive mating than the other type. Even for two women with the same education background and even everything else is the same, but if they have different incentives in their education investment decisions, their marriage prospects could be different. The women who marrying for money will have a higher probability to find a highly educated husband than the women who marrying for romance, given they have the same education level. Also it is worth noting that the estimation value of the women's education in table A2.5 is smaller than that in the type model for the type of marrying for money and larger than that in the type model for the type of marrying for romance.

But what characteristics can be used to identify the type? Theoretically, according to my knowledge, there is no such a theory to help me to identify the motivation behind the female's education investment decision. In the chapter, it is assumed that the determination of the type is correlated with the women's attitude about the marriage, whether they are religious or where they are from.

Table A2.7 reports the type determination results. I include the region variable, the family background variable and the women's attitude to the women's role in the family in the model. It turns out that the women's attitude about the female's role in the family is the key variable to determine the women's type. The results suggest that women with a traditional attitude about the women's role in the family are more likely to be the type of marrying for money, i.e. their education investment is at least partly driven by their marriage prospects. The women who think it is a better arrangement if women take care of the family usually value the family more than the career and their marriage prospect is more likely to be an important consideration behind their education investment. The variable whether females stay with their parents at the age 14 has a positive but insignificant estimate in the type equation. It suggests that the females who are more dependent on their parents have a higher probability to be the type of marrying for money. Other variables do not have significant impacts on the determination of the type.

Based on the estimation results from table A2.7, table A2.8 reports the distribution of women's education level. Panel A lists women's average schooling years when getting married by the type and the husband-education status. Overall, the type of marrying for romance has more years of schooling than that of the type of marrying for money. The average education for the marry-for-money type is 12.2 years, while it is 13.1 years for the type of marry-for-romance. This result is contrary to what Booth and Coles (2005) found. According to the results from the type determination, the marry-for-money type women usually have the traditional attitude about the women's status in the marriage. There is a high probability for them to stay home after getting married. Then it is costly for them to make too much investment on education. However, on the average, the education level for the type of marry for money is still above 12 years, which implies most of women who marry for money just have the college education experience, but most of they do not finish the college education. And it is also worth noting that women who marry a good husband on the average have higher education than those who do not regardless of the type.

Panel A also reports the observation number for each type. Among the 2,342 women, 61% (or 1,431) belong to the marrying for romance type and 39% (or 911) belong to the marrying for money type. Among the type of marrying for romance, 40% of them find good husbands and among the type of marrying for money, 32% of them find good husbands.

Panel B presents women's average schooling years at the time of marriage across the type and their attitude about the marriage. As we can see, the women with the traditional attitude about the marriage have less education than those with non-traditional attitude. Because the attitude about the marriage is the most significant factor to determine the women's type, it is not surprising that the type of marrying for romance has lower education than the type of marrying for money. Panel B also has the information of the number of observations. Almost 82% of non-traditional women are the type of marrying for romance, and around 76% of traditional women are the type of marrying for money.

Panel C is the distribution of the education across the marriage results and the attitude. Around 41% of the women who do not have the traditional attitude about the marriage find a good husband, while only around 31% of traditional women find a good husband. From Panel C and D, we can see that 54% of non-traditional women go to college, while only 41% of them find good husbands. Among the traditional women, about 40% of them go to college and 31% of them find good husbands. Also among the traditional women, 60% of them do not go to school and 69% of them do not find good husbands. Among non-traditional women, 46% of them do not have college education and 59% of them do not find highly-educated husbands.

3.6 Conclusion

To explain the rapid increase in the women's educational attainment, this chapter investigates two main motivations of investing the education. One is the marriage market return, the other is the job market return. Different motivations behind the educational investment could result in different investment behaviors and marriage prospects. According to different motivations, this chapter assumes that there are two types of women: one is the type of marrying for money whose education investment is at least partly driven by their marriage market return, and the other is the type of marrying for romance whose education investment has no marriage considerations. The results give strong evidences of the existence of the two types of women. It also shows that the positive assortive mating in the education is more significant for the type of marrying for money than that for the type of marrying for romance. It implies that different incentives for the education will results in the different assortive mating in the education. Also this chapter finds that although the education will help the type of marrying for money more in finding good husbands, on the average their educational attainment is lower than the type of marrying for romance. Furthermore, this chapter also finds that the women's attitude about the women's status in the marriage is the main factor to determine their motivation for education. A woman with more traditional ideas about the women's status in the marriage is more likely to go to college to find a good husband. In the last several decades, the women's socioeconomic status has been greatly improved. More and more women get out pursuing their own career. Such a radical change in the women's status in the marriage leads to more marry-for-romance type women and therefore a rapid increase in the women's education attainment.

CHAPTER IV

CONCLUSION

This dissertation makes three main contributions to the neighborhood effect literature. First, it contributes to the econometric solution to the neighborhood effect by proposing a new model to identify the neighborhood effect. The key feature of the proposed model is its treatment of the unobserved types. The dissertation presents two reasons to justify its assumption of unobserved types. The first reason is more theoretical. The dissertation argues that people move for different reasons. One reason is to search for neighborhoods that may help education. In this case, the moving decision is endogenous. However, it is also typical for people to move for reasons that are independent to education choice. In this case, the moving decision is potentially exogenous. The second reason is empirical. The test based on Heckman sample selection model shows that coefficients for inverse mills ratio are different between the mover group and the non-mover group.

Second, this dissertation provides strong evidences of the existence of the two unobserved types. The estimation is conducted by allowing possibilities of both the one-type model and the two-type model. The data supports the two-type model. In addition, people from different types exhibit different behaviors. The endogenous type has 91.7% more schooling years (beyond the compulsory schooling) than the exogenous type.

Third, introducing the two unobserved types of individuals helps us to identify the neighborhood effect. When I use the Heckman sample selection model, I do not find any statistically significant neighborhood effects. However, the type-consistent model shows that the neighborhood effect only shows up for the endogenous mover group. Therefore, it is not surprising that the effect is not seen for the population with one-type model. I find that neighborhood difference between blacks and whites is an important factor affecting their educational gap. In particular, I find that the county-level race specific percentage of high school graduates can explain 37.7% of education difference between blacks and whites.

To explain the rapid increase in the women's educational attainment, this chapter investigates two main motivations of investing the education. One is the marriage market return, the other is the job market return. Different motivations behind the educational investment could result in different investment behaviors and marriage prospects. According to different motivations, this chapter assumes that there are two types of women: one is the type of marrying for money whose education investment is at least partly driven by their marriage market return, and the other is the type of marrying for romance whose education investment has no marriage considerations. The results give strong evidences of the existence of the two types of women. It also shows that the positive assortive mating in the education is more significant for the type of marrying for money than that for the type of marrying for romance. It implies that different incentives for the education will results in the different assortive mating in the education. Also this chapter finds that although the education will help the type of marrying for money more in finding good husbands, on the average their educational attainment is lower than the type of marrying for romance. Furthermore, this chapter also finds that the women's attitude about the women's status in the marriage is the main factor to determine their motivation for education. A woman with more traditional ideas about the women's status in the marriage is more likely to go to college to find a good husband. In the last several decades, the women's socioeconomic status has been greatly improved. More and more women get out pursuing their own career. Such a radical change in the women's status in the marriage leads to more marry-for-romance type women and therefore a rapid increase in the women's education attainment.

REFERENCES

Aaronson, D. (1998). “Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes.” *Journal of Human Resources*, 33 (4):915-946.

Ananat, O. Elizabeth (2007). “The Wrong Side(s) of the Tracks Estimating the Causal Effects of Racial Segregation on City Outcomes.” *National Bureau of Economic Research Working Paper*, No.13343.

Becker, G. S. and Richard Posner (2008). “The New Gender Gap in Education-Becker.” http://www.becker-posner-blog.com/archives/2008/03/the_new_gender_1.html, The Becker-Posner Blog.

Behrman, J. R. and M. R. Rosenzweig (2002). "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?" *The American Economic Review* 92(1): 323-334.

Behrman, J. R., M. R. Rosenzweig, and P. Taubman (1994). "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment." *The Journal of Political Economy* 102(6): 1131-1174.

Benham, L. (1974). "Benefits of Women's Education within Marriage." The *Journal of Political Economy* 82(2): S57-S71.

Booth, A. L. and M. Coles (2005). "Education, Matching and the Allocative Value of Romance." *The Center for Economic Policy Research Discussion Paper* No. 5099.

Borjas, G., S. Bronars, and S. Trejo (1992). "Self-Selection and Internal Migration in the United States." *Journal of Urban Economics*, 32:159-185.

Boulier, B. L. and M. R. Rosenzweig (1984). "Schooling, Search, and Spouse Selection: Testing Economic Theories of Marriage and Household Behavior." The *Journal* of Political Economy 92(4): 712-732.

Bowles, Samuel (1970). "Migration as Investment: Empirical Tests of the Human Investment Approach to Geographical Mobility." *Review of Economics and Statistics*, 52(4): 356-362.

Brock, W.A. and Steven N. Durlauf (2001). "Discrete Choice with Social Interactions." *Review of Economic Studies*, 68:235-260.

Brooks-Gunn, J., G. J. Duncan, and J. L.e. Aber (1997). *Neighborhood Poverty I: Context and Consequences for Children*. New York, Russell Sage.

Cancian, M., Sheldon Danziger, Peter Gottschalk. (1993). "Working Wives and Family Income Inequality among Married Couples. In Uneven tides: Rising inequality in America." the Annual Meeting of the Population Association of America, Denver, Colorado, April 30 - May 2, 1992.

Card, D., A. Mas, and J. Rothstein, (2007). "Tipping and the Dynamics of Segregation in Neighborhoods and Schools." *National Bureau of Economic Research Working Paper* No.13052.

Chen, Y. and S. Rosenthal (2006). "Local Amenities and Life Cycle Migration: Do People Move for Jobs or Fun." *Journal of Urban Economics*, 64(3):519-537.

Coleman, J.S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York (1979). *Equality of Educational Opportunity*. New York, Arno Press.

Costa, D. and M. Kahn (2000). "Power Couples: Changes in the Locational Choice of the College Educated, 1940-1990." *Quarterly Journal of Economics*, 115 (4):1287-1315.

Couch, K. and Mary C. Daly (2002). "Black-White Wage Inequality in the 1990s: A Decade of Progress." *Economic Inquiry*, 40 (1):31-41.

Crane, J. (1991). "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping out and Teenage Childbearing." *The American Journal of Sociology*, 96 (5):1226-1259.

Cutler, David M., Edward L. Glaeser, Jacob L.Vigdor, 1999. "The Rise and Decline of the American Ghetto." *Journal of Political Economy*, 107 (June 1999): 455-506.

Datcher, L. (1982). "Effects of Community and Family Backgrounds on Achievement." *Review of Economics and Statistics*, 64:32-41.

Duncan, G. (1994). "Families and Neighbors as Sources of Disadvantage in the Schooling Decision of White and Black Adolescents." *American Journal of Education*, 1003:20-53.

Duncan, G., James R. Connell, and Pamela Klebanov (1997). "Conceptual and Methodological Issues in Estimating Causal Effects of Neighborhoods and Family Conditions on Individual Development." In *Neighborhood Poverty:*

Context and Consequences for Children, vol I, ed. J. Brooks-Gunn, G. Duncan, and L. Aber. New York: Russell Sage Foundation.

Durlauf, S.N. (2002). “On the Empirics of Social Capital.” *Economic Journal,* 112 (483):459-479.

Feinstein, J. S. (1990). “Detection Controlled Estimation.” *Journal of Law and Economics* 33(1): 233-276.

Fernández, R., Alessandra Fogli (2002). “Marrying Your Mom: Preference Transmission and Women’s Labor and Education Choices.” *National Bureau of Economic Research Working Paper,* No.W9234.

Gan, Li, and Roberto Mosquera. “An Empirical Study of the Credit Market with Unoserved Consumer Types.” *National Bureau of Economic Research Working Paper*, No.13873.

Gibbons, S. (2002). “Neighborhood Effects on Educational Achievement: Evidence from the Census and National Child Development Survey.” *Centre for the Economics of Education Working paper No. Dp 18*.

Ginther, D., Haveman,R.,and Wolfe,B. (2000). “Neighborhood Attributes as Determinants of Children's Outcomes: How Robust are the Relationships?” *Journal of Human Resources*, 35 (4):603-642.

Grossbard-Shechtman, S., Shoshana Neuman., Ed. (1991). “Cross-Productivity Effects of Education and Origin on Earnings: Are They Really Reflecting Productivity?” In *Handbook of Behavioral Economics*. Greenwich, Connecticut, JAI Press.

Hanratty, M., Sara S. McLanahan and Becky Pettit (1998). “The Impact of the Los Angeles Moving to Opportunity Program on Residential Mobility, Neighborhood Characteristics, and Early Child and Parent Outcomes.” *Working Paper*, 98-18.

Hastings, J.S., Thomas J. Kane and Douglas O. Staiger, (2006). “Preferences and Heterogeneous Treatment Effects in a Public School Choice Lottery.” *National Bureau of Economic Research Working Paper* No.12145.

Ioannides, Y.M. and Giulio Zanella (2007). “Searching for the Best Neighborhood: Mobility and Social Interactions.” *Department of Economics University of Siena Working Paper* No.533.

Ioannides, Y.M. and Jeffrey E. Zabel (2008). “Interactions, Neighborhood Selection, and Housing Demand.” *Journal of Urban Economics*, 63: 229-252.

Jacob, B., (2004). “Public Housing, Housing Vouchers and Student Achievement: Evidence from Public Housing Demolitions in Chicago.” *American Economic Review*, 94 (1):233-258.

Jepsen, L. K. (2005). “The Relationship between Wife’s Education and Husband’s Earnings: Evidence from 1960 to 2000.” *Review of Economics of the Household* 3(2): 197-214.

Juhn, C. and K. M. Murphy (1997). “Wage Inequality and Family Labor Supply.” *Journal of Labor Economics* 15(1): 72-97.

Katz, L.F., Kling, J.R., and Liebman, J. B, (2001). “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment.” *Quarterly Journal of Economics*, 116 (2):607-654.

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

Kennan, J., and J. Walker (2006). "The Effect of Expected Income on Individual Migration Decisions." *National Bureau of Economic Research Working Paper*, No.9585.

Kling, J., and Mark Votruba, (2001). "Mobility of Families in the Gautreaux Housing Assistance Program." *Princeton University Working Paper*, No.243.

Knittel, Christopher, and Victor Stango (2003). "Price Ceilings as Focal Points for Tacit Collusion: Evidence from Credit Cards." *American Economic Review*, 93 (5) (Dec., 2003): pp 1703-1729.

Ladd, H.F., and Jens Ludwig, (1997). "Federal Housing Assistance: Residential Relocation, and Educational Opportunities: Evidence from Baltimore." *American Economic Review*, 87:272-277.

Lam, D. and R. F. Schoeni (1994). "Family Ties and Labor Markets in the United States and Brazil." *Journal of Human Resources*, 29: 1235-1258.

Lillard, D., (1993). "Neighborhood Effects on Educational Attainment." *CEH Working Paper RP93-5*, Cornell University, Ithaca, NY.

Lleras-Muney, A., (2005). "The Relationship between Education and Adult Mortality in the United States." *National Bureau of Economic Research Working Paper* No.8986.

Ludwig, J., Helen F. Ladd, and Greg J. Duncan, (2001). "Urban Poverty and Educational Outcomes." *Brookings-Wharton Papers on Urban Affairs* (147-201).

Manski, C.F., (1993). “Identification of Endogenous Social Effects: The Reflection Problem.” *Review of Economic Studies*, 60 (3):531-542.

Mare, R. D. (1991). “Five Decades of Educational Assortive Mating.” *American Sociological Review* 56(1): 15-32.

Moffitt, R., (2001). “Policy Interventions, Low-Level Equilibria and Social Interactions.” *Social Dynamics*, Steven N. Durlauf and H. Peyton Young. Cambridge, Mass.: MIT Press.

Neal, D.A. and W.R. Johnson (1996) “The Role of Premarket Factors in Black-White Wage Differences.” *The Journal of Political Economy*, 104 (5): 869-895.

O’Neill J.E. (1990) “The Role of Human Capital in Earnings Differences between Black and White Men.” *Journal Economic Perspectives,* 4 (Fall 1990): 25-45.

Oreopoulos, P., (2003). “The Long-Run Consequences of Living in a Poor Neighborhood.” *Quarterly Journal of Economics*, 118 (4):1533-1575.

Plotnick, R., and Saul Hoffman (1999). “The Effect of Neighborhood Characteristics on Young Adult Outcomes: Alternative Estimates.” *Social Science Quarterly*, 80 (1):1-18.

Popkin, S.J., James E. Rosenbaum, and Patrica M. Meaden (1993). “Labor Market Experiences of Low-Income Black Women in Middle-Class Suburbs: Evidence from a Survey of Gautreaux Program Participants.” *Journal of Policy Analysis and Management*, 12:556-573.

Rivers, D. and Q.H.Vuong, (1988). “Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models.” *Journal of Econometrics*, 39:347-366.

Schwartz, C. R., Robert D.Mare (2005). “Trends in Educational Assortive Marriage from 1940 to 2003.” *Demography*, 42:621-646.

Scully, G. W. (1979). “Mullahs, Muslims, and Marital Sorting.” *The Journal of Political Economy* 87(5): 1139-1143.

Weinberg, Daniel H. and John Iceland “Racial and Ethnic Residential Segregation in the United States: 1980-2000.” *U.S. Census Bureau 2000 Special Reports*, CENSR-3

Weiss, Y. and R. J. Willis (1997). “Match Quality, New Information, and Marital Dissolution.” *Journal of Labor Economics* 15(1): S293-S329.

Weitzman, Martin. 1979. “Optimal Search for the Best Alternative.” *Econometrica,* 47(3): 641—654.

Wong, Y.-C. (1986). “Entrepreneurship, Marriage, and Earnings.” *The Review of Economics and Statistics* 68(4): 693-699.

APPENDIX A

Table A1.1-1 Data description---individual variables

Variables	White	Black	All
Schooling years beyond compulsory schooling	5.3149	4.9301	5.1880
	(2.7780)	(2.3929)	(2.6755)
Move dummy (1 if moved in 1979 and 1982)	0.1682	0.1244	0.1533
	(0.3741)	(0.3302)	(0.3603)
Age	15.61	15.60	15.60
	(1.06)	(1.06)	(1.06)
Mom highest grade (total years of education)	10.92	10.75	10.73
	(3.08)	(2.46)	(3.05)
Dad highest grade (total years of education)	11.23	10.23	10.83
	(3.66)	(2.90)	(3.54)
The oldest sibling's highest grade	11.85	11.81	11.81
	(2.52)	(3.08)	(3.01)
Family size	4.96	5.62	5.19
	(1.78)	(2.31)	(2.01)
Family income (in $)	18,671	10,876	16,087
	(12,243)	(8,585)	(11,665)
Safe feeling about the school (1 if feel safe)	0.5924	0.5368	0.5724
	(0.4915)	(0.4989)	(0.4948)
Religion Degree	0.4039	0.4201	0.4102
	(0.4908)	(0.4938)	(0.4919)
Urban (1 if living in urban)	0.74	0.7856	0.7653
	(0.44370)	(0.4106)	(0.4239)
Family poverty (1 if in poverty)	0.2239	0.5005	0.3125
	(0.4169)	(0.5002)	(0.4636)
Living with parents until age 18 (1 if yes)	0.6270	0.4584	0.5754
	(0.4837)	(0.4985)	(0.4943)
Multi move in 1979 and 1982 (1 if yes)	0.0707	0.0507	0.0638
	(0.2564)	(0.2195)	(0.2443)
Living with parents at age 14 (1 if yes)	0.7217	0.4469	0.6381
	(0.4483)	(0.4974)	(0.4806)
Married (1 if never married)	0.9825	0.9971	0.9872
	(0.1311)	(0.0535)	(0.1122)
Black (1 if black)	0	1	0.2835
	0	0	(0.4508)
Male (1 if male)	0.4835	0.5005	0.4891
	(0.4998)	(0.5002)	(0.5000)

Table A1.1-2 Data description---county-level variables

Variables	White	Black	All
Percentage high school graduates	51.01	45.15	49.39
	(10.68)	(10.97)	(10.97)
Percentage high school graduates (race specific)	53.68	32.13	46.21
	(11.36)	(10.61)	(15.33)
Crime rate (per 100,000)	4,949	5,833	5395
	(3154)	(3769)	(3693)
Mortality rate (per 100,000)	89.21	161.98	109.78
	(69.00)	(150.34)	(105.50)
Percentage employed in the education institution	7.97	7.33	7.82
	(2.70)	(1.78)	(2.53)
Per capita income (in $)	4,362	4,293	4,345
	(873)	(954)	(896)
Percentage of population is black	8.900	25.40	13.67
	(10.05)	(14.24)	(13.55)
Civil unemployment rate	4.51	4.63	4.61
	(1.75)	(1.92)	(1.82)
Percentage of population being urban	69.29	74.07	71.41
	(27.60)	(29.67)	(28.06)
Number of observations	2,461	1,045	3,686

Table A1.2 Percentage difference between movers and non-movers

	Blacks		Whites	
	Before move % diff	**After move % diff**	**Before move % diff**	**After move % diff**
Mortality rate	-1.1298	-10.8796	2.3601	-6.7735
Percentage employed rate in education	2.8601	27.8552	3.6462	25.5653
Per capita income	-1.9702	0.7241	-0.5589	-0.7654
Percentage of black	-2.3586	-16.0754	-8.4888	-3.5056
Crime rate	-4.1961	-0.0259	-5.6378	3.7392
Percentage of high school graduates	-0.9830	9.3583	0.4812	3.4753
Percentage of urban	-6.8416	-1.2992	-6.4887	-2.2933
Civil unemployment rate	2.3718	-9.4775	-1.4453	-0.9936

Note: (1) Because 1982 data are not available in IPUMS, I could not impute county level race-specific percentage of high school graduates here.
(2) A negative number in the table indicates that the county variable for movers is less than that for the non-movers.

Table A1.3 OLS and Heckman estimation for movers and non-movers

Variables	Movers OLS	Movers Heckman	Non-Movers OLS	Non-Movers Heckman
Religion degree	1.0080***	1.0170***	0.6000***	0.5900***
	(0.2350)	(0.2350)	(0.0851)	(0.0850)
Family poverty	-1.3220***	-1.2890***	-0.7700	-0.8030***
	(0.2860)	(0.2850)	(0.0966)	(0.0966)
Mother highest grade	0.2860***	0.2950***	0.1920***	0.1930***
	(0.0394)	(0.0395)	(0.0148)	(0.0148)
Predicted percentage high school	0.0080	0.0080	0.0104*	0.0080
graduates(race specific)	(0.0160)	(0.0160)	(0.0050)	(0.0050)
Mortality rate	0.0012	0.0012	0.0003	-0.0001
	(0.0012)	(0.0012)	(0.0004)	(0.0004)
Crime rate	-4.30e-05	-4.33e-05	-4.93e-06	-7.23e-06
	(3.43e-05)	(3.40e-05)	(1.54e-05)	(1.55e-05)
Civil unemployment rate	0.091	0.102	0.209***	0.222***
	(0.066)	(0.066)	(0.026)	(0.026)
Percentage employed in education	0.035	0.028	0.033	0.022
	(0.043)	(0.043)	(0.020)	(0.020)
Percentage of being urban	0.007	0.008	0.002	0.003
	(0.006)	(0.006)	(0.002)	(0.002)
Percentage of being black	.0219*	.0213*	0.008**	0.009**
	(0.012)	(0.012)	(0.004)	(0.004)
Per capita income	0.0003	0.0003	8.17e-05	4.85e-05
	(0.0002)	(0.0002)	(8.72e-05)	(8.77e-05)
Constant	(0.3790)	(0.4120)	0.6550	0.867**
	(1.0850)	(1.0850)	(0.4170)	(0.4190)
Inverse Mills Ratio		0.6040		2.0460***
		(1.2850)		(0.7730)
Observations	604	615	3398	3467
R^2	0.2090	.0.2090	0.1290	.0.1292

Note: (1) The estimation includes dummy variables for missing values of imputed variables.

(2) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table A1.4 IV Estimation for the county high school graduates

Variables	**Coefficient**	**Std Err**
County mean inverse mills ratio	-0.7217***	0.2511
Absolute Centralization Index (ACE)	-9.6376	6.7710
Absolute Clustering Index (ACL)	-25.6396***	5.6340
Atkinson Index with b=#1 (A1)	-11.7859	19.4012
Entropy Index (H)	87.1056***	19.4236
Gini Index (G)	-44.6909***	9.8105
Relative Centralization Index (RCE)	5.3356	5.2515
Relative Concentration Index (RCO)	-13.1328**	5.2579
Inverse mills ratio	0.8593***	0.2883
Religion degree	-0.5298**	0.2536
Family poverty	-0.0986	0.4005
Mom highest grade	0.3348***	0.0709
Mortality rate	0.0026	0.0029
Crime rate	-0.0002	0.0001
Civil unemployment rate	0.3859	0.3149
Employment in education institution	1.0655***	0.1803
Percentage being urban	0.0744***	0.0258
Percentage being black	-0.1806***	0.0417
Per capita income	0.0080***	0.0008
Constant	2.7008	4.4166
Observations	4100	
R^2	0.851	

Note: (1) The estimation includes dummy variables for missing values of imputed variables.
(2) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table A1.5 Results for IV probit for the moving decision

	Variables	**IV Probit**	**Std Err**
Individual variables	Mom's highest grade	0.0311***	0.0108
	Dad's highest grade	0.0193**	0.0088
	Family size	-0.0436***	0.0133
	Age	2.8030***	0.8050
	Age squared	-0.0844***	0.0257
	Black (1 if black)	-0.0975	0.0659
	Male (1 if male)	-0.0442	0.0488
	Urban (1 if living in Urban)	-0.0162	0.1210
The destination	Percentage urban	0.025***	0.005
	Percentage female head	0.021	0.031
	Crime rate	6.81e-05***	2.45e-05
	Percentage high school graduates	0.0550***	0.0120
	Percentage college graduates	0.0888***	0.014
	Civil unemployment rate*	-0.353***	0.137
	Civil employment	-1.32e-07	1.57e-07
	Per capita income	-0.0013***	0.0004
The origin	Percentage urban	-0.026***	0.006
	Percentage female head	-0.034	0.035
	Crime rate	-5.28e-05**	2.45e-05
	Percentage high school graduates	-0.0610***	0.0130
	Percentage college graduates	-0.028	0.017
	Civil unemployment rate*	0.354***	0.135
	Civil employment	-1.32e-06	1.57e-06
	Per capita income*	0.0013***	0.0004
	Constant	-24.5800***	6.29
Observations		4,817	

Note: (1) This is the estimation results for the equation (2) in the paper.
(2) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
(3) Variables with superscript * are endogenous and the predicted values used here.
(4) Instrument variables for those variables are the mean values of adjacent counties excluding the residence
(5) The estimation also includes dummy variables for missing values of imputed variables.

Table A1.6 Second stage results for Rivers and Vuong procedure

	Variables	Estimation
Destination	Error for percentage urban	-9.13E-06
		(0.0016)
	Error for percentage female head	-0.0024
		(0.0111)
	Error for the crime rate	-1.01e-05
		(0.0001)
	Error for percentage high school graduates	0.0028
		(0.0035)
	Error for percentage college graduates	-0.0097
		(0.0091)
	Error for unemployment rate	0.0198**
		(0.0098)
	Error for civil employment	-9.72E-07
		-6.81E-07
	Error for per capita income	0.0003
		-0.0003
Origin	Error for percentage urban	0.0020
		-0.0019
	Error for percentage female head	0.0100
		-0.0112
	Error for the crime rate	-7.06E-05
		-9.25E-05
	Error for percentage high school graduates	0.0027
		-0.0034
	Error for percentage college graduates	0.0142
		-0.0099
	Error for unemployment rate	-0.0168*
		-0.0091
	Error for civil employment	4.78E-07
		-8.07E-07
	Error for per capita income	-0.0006**
		-0.0003

Note: (1) This procedure is to justify the endogeneity of variables in the equation (2).
(2) Only the estimations for the error terms presented here
(3) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table A1.7-1 Type consistent model regression results for type determination

Variables	Coefficient	Std Err
Constant	-0.3448	2.0360
Family size	0.0266**	0.0111
Living with parents until age 18	-0.0510*	0.0289
Black (1 if black)	0.0097	0.0431
Mom highest grade	-0.0517**	0.0233
Male (1 if male)	0.1449**	0.0597
Urban (1 if living in urban)	0.1121	0.0858
Move multiple times (1 yes)	-0.0339	0.0384
Family poverty	-0.0175	0.0991
the oldest sibling highest grade	-0.0335**	0.0133
Dad highest grade	-0.0297***	0.0111
Living with parents at age 14(1 yes)	-0.0125	0.0252
Married (1 if never married)	-0.1944*	0.1182
Family income (in US $)	-5.55E-06**	2.48E-06
Age	0.1864	0.2671
Age squared	-0.0063	0.0086
Feel safe about the school (1 if yes)	-0.1228***	0.0426

Note: (1) The estimation includes dummy variables for missing values of imputed variables.
(2) Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A1.7-2 Type consistent model regression results for school equation

Variables	Exogenous Coefficient	Endogenous Coefficient.
zeta	0.6825	-0.6661^{*}
	(0.7439)	(0.3801)
Predicted percentage high school graduates (race specific)	-0.0480	0.0290
	(0.0370)	(0.0240)
Predicted percentage high school graduates (race specific)*move	-0.0470	0.0510^{***}
	(0.0360)	(0.0140)
Religion degree	-0.0661	0.8676^{***}
	(0.4753)	(0.2720)
Family poverty	-0.1886	-0.2219
	(0.5301)	(0.6509)
Mom highest grade	-0.5071^{**}	0.1798
	(0.2359)	(0.1162)
Mortality rate	-0.0022	0.0056^{***}
	(0.0027)	(0.0019)
Crime rate	-4.58E-05	2.94E-05
	8.77E-05	5.51E-05
Civil Unemployment rate	0.249^{**}	0.116
	(0.111)	(0.100)
Employment in education institution	0.189	-0.043
	(0.117)	(0.076)
Percentage urban	0.003	0.013
	(0.014)	(0.010)
Percentage black	-0.025	0.033^{*}
	(0.027)	(0.018)
Per capita income	0.0007	-0.0003
	(0.0005)	(0.0003)
Constant	1.0830	4.0591^{*}
	(2.0362)	(2.3273)

Note: (1) The estimation includes dummy variables for missing values of imputed variables.
(2) Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A1.8 Education level across the type and the move

Type	Non-movers	Movers	Total
Endogenous type	5.3226 (2.5099)	6.5143 (2.9996)	5.5147 (2.6314)
# Observations	2,539	488	3027
Exogenous type	3.7234 (2.3382)	3.4156 (2.3971)	3.6874 (2.3454)
# Observations	582	77	659
Total	5.0244 (2.5555)	6.0920 (3.1106)	5.1880 (2.6755)
# Observations	3121	565	3686

Note: the number in the bracket is the standard deviation

Table A1.9 Observed and predicated education level by type and race

Type	White	Black	Other	Total
Panel A: Observed schooling years				
Endogenous type	5.6370 (1.3295)	5.2437 (1.0738)	5.6989 (1.1049)	5.5285 (1.2689)
Exogenous type	3.5012 (0.9679)	3.5511 (1.2166)	4.2751 (0.9010)	3.6241 (0.9658)
Total	5.3150 (1.4920)	4.9262 (1.2308)	4.9712 (1.2310)	5.1880 (1.4218)
Panel B: Predicted schooling years				
Endogenous type	5.6311 (2.7203)	5.2026 (2.3735)	5.7614 (2.6261)	5.5147 (2.6314)
Exogenous type	3.5337 (2.4051)	3.7500 (2.1059)	4.1739 (2.5316)	3.6874 (2.3454)
Total	5.3149 (2.7780)	4.9301 (2.3929)	4.9500 (2.6913)	5.1880 (2.6755)

Note: the number in the bracket is the standard deviation

Table A2.1 Data description

Variable	Mean	Std. Dev.	Min	Max
Father's highest grade	10.86635	3.956468	0	20
Mother's highest grade	10.88471	3.1707	0	19
Family poverty status	0.230572	0.421289	0	1
Urban or not (1 is from urban)	0.757899	0.428447	0	1
Unemployment rate by county	2.568745	0.740144	1	5
Age when getting married	23.50769	4.222727	15	41
Live with parents at age 14	0.775833	0.417122	0	1
Religious (1 is yes)	0.364646	0.481433	0	1
Tradition attitude to the women's status (1 is yes)	0.362084	0.480706	0	1
Living in the north east	0.166951	0.373012	0	1
Living in the north central	0.274125	0.446167	0	1
Living in the south	0.373185	0.483754	0	1
Wife's highest grade when married	12.72417	2.389087	0	20
Wife's family poverty status when married	0.102904	0.303898	0	1
Wife's residence is urban or not when married (1 is urban)	0.771136	0.420191	0	1
Educ (1 if going to college)	0.490606	0.500019	0	1
Marriage (1 if husband's education level is relatively high)	0.374466	0.484088	0	1
Observations	2342			

Table A2.2 Marriage prospect distribution across female's education level

College\Marriage	0	1	Total
0	1,002	191	1,193
1	463	686	1,149
Total	1,465	877	2,342

Table A2.3 Probit estimation for college equation

Variables	Probit (Marginal effects reported)	Probit
Mother's highest education	0.0842***	0.211***
	(0.0147)	(0.0369)
Dad's highest education	0.128***	0.320***
	(0.0148)	(0.0370)
Residence is urban or not	0.0570***	0.143***
	(0.0110)	(0.0277)
Family poverty status	-0.0181	-0.0454
	(0.0112)	(0.0281)
Unemployment rate	-0.00288	-0.00721
	(0.0109)	(0.0274)
Observations	2342	2342

Note: (1) Standard errors in parentheses
(2) *** p<0.01, ** p<0.05, * p<0.1
(3) All of the variables are standardized

Table A2.4 Probit estimation for marriage equation

Variables	Probit (Marginal effects reported)	Probit
Female's highest grade when married	0.232***	0.632***
	(0.0158)	(0.0429)
Mother's highest education	0.0458***	0.125***
	(0.0126)	(0.0343)
Residence is urban or not when married	0.0623***	0.170***
	(0.0115)	(0.0315)
Family poverty status when married	-0.00214	-0.00582
	(0.0125)	(0.0341)
Female's age when married	0.334***	0.909***
	(0.0993)	(0.271)
Female's age squared	-0.289***	-0.786***
	(0.0960)	(0.262)
Observations	2342	2342

Note: (1) Standard errors in parentheses
(2) *** p<0.01, ** p<0.05, * p<0.1
(3) All of variables are standardized

Table A2.5 Estimation results for the type of marrying for money only

	Variables	MLE	Probit
College Equation	Mother's highest education	0.226724***	0.211***
		(0.025638)	(0.0369)
	Dad's highest education	0.298634***	0.320***
		(0.032926)	(0.0370)
	Residence is urban or not	0.151182***	0.143***
		(0.021869)	(0.0277)
	Family poverty status	-0.0469	-0.0454
		(0.026353)	(0.0281)
	Unemployment rate	-0.02399	-0.00721
		(0.016099)	(0.0274)
Marriage Equation	Female's highest grade when married	0.718869***	0.632***
		(0.253956)	(0.0429)
	Mother's highest education	0.256588***	0.125***
		(0.070498)	(0.0343)
	Residence is urban or not when married	0.248346***	0.170***
		(0.081112)	(0.0315)
	Family poverty status when married	-0.02236	-0.00582
		(0.047996)	(0.0341)
	Female's age when married	0.438431	0.909***
		(0.395786)	(0.271)
	Female's age squared	-0.28368	-0.786***
		(0.393775)	(0.262)
	Lambda	0.329833***	
		(0.032087)	
	Observations	2342	2342

Note: (1) Standard errors in parentheses
(2) *** p<0.01, ** p<0.05, * p<0.1
(3) All of variables are standardized

Table A2.6 Estimation results for the type model

	Variables	Marry for money estimates	Marry for romance estimates
College Equation	constant	-1.60085***	0.541267
		(0.044134)	(0.318006)
	Mother's highest education	1.32763***	0.20695***
		(0.01734)	(0.074773)
	Dad's highest education	1.68447***	0.079347
		(0.001578)	(0.112578)
	Residence is urban or not	-0.06736	0.264757
		(0.333753)	(0.195194)
	Family poverty status	0.00464	-0.10232**
		(0.33552)	(0.054679)
	Unemployment rate	-0.25579***	0.121511
		(0.005847)	(0.17848)
Marriage Equation	constant	-0.0303	-0.2499
		(0.282709)	(0.064054)
	Female's highest grade when married	1.394834***	0.319246***
		(0.149968)	(0.132329)
	Mother's highest education	0.200312	0.003884
		(0.948854)	(0.072968)
	Residence is urban or not when married	0.047217	0.551781
		(0.393309)	(0.160666)
	Family poverty status when married	-0.01961	-0.01443
		(0.288581)	(0.161984)
	Female's age when married	0.314433*	0.569898
		(0.22338)	(0.382333)
	Female's age squared	-0.23397	-0.46441
		(0.474527)	(0.355544)
	sigma	0.321755***	
		(0.0328623)	

Note: (1) Standard errors in parentheses
(2) *** p<0.01, ** p<0.05, * p<0.1
(3) All of variables are standardized

Table A2.7 Type model estimation

Variables	Estimates
constant	-0.0821
	(0.3684)
Living with parents at age 14	0.0171
	(0.1070)
Religious	-0.2123
	(0.1992)
Tradition attitude to the women's status	0.2819***
	(0.0884)
Living in the north east	-0.0881
	(0.1386)
Living in the north central	0.1719
	(0.2184)
Living in the south	-0.0933
	(0.1705)

Note: (1) Standard errors in parentheses
(2) *** p<0.01, ** p<0.05, * p<0.1

Table A2.8 the distribution of Women's education level when getting married

Panel A

Type\Marriage		0	1	Total
Marry for romance	Women's education level	12.03077	14.53584	13.0566
	#observation	845	586	1431
Marry for money	Women's education level	11.5	13.6976	12.20198
	#observation	620	291	911
Total	Women's education level	11.80614	14.2577	12.72417
	#observation	1465	877	2342

Panel B

Type\Attitude to the marriage		0	1	Total
Marry for romance	Women's education level	13.125	12.65217	13.0566
	#observation	1224	207	1431
Mary for money	Women's education level	12.61852	12.02652	12.20198
	#observation	270	641	911
Total	Women's education level	13.03347	12.17925	12.72417
	#observation	1494	848	2342

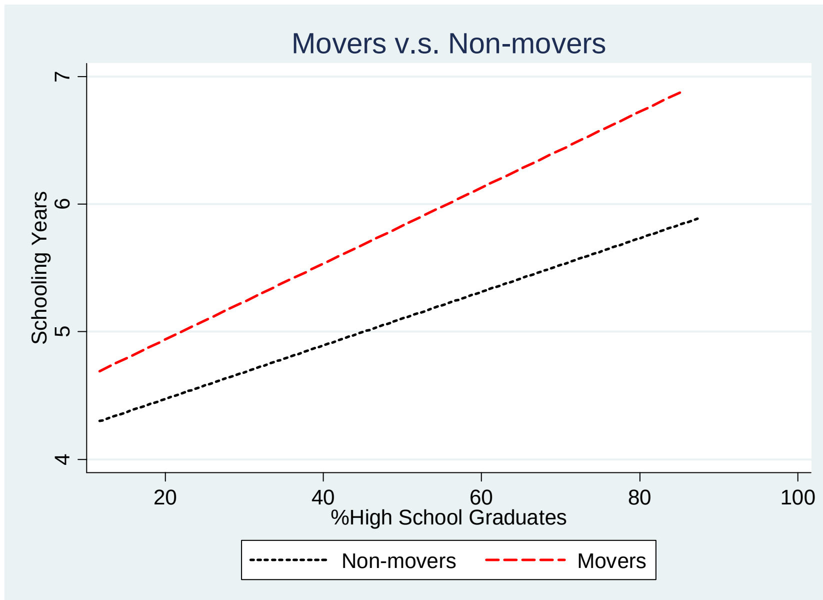
Panel C

Marriage\Attitude to the marriage		0	1	Total
0	Women's education level	12.0703	11.40652	11.80614
	#observation	882	583	877
1	Women's education level	14.42157	13.87925	14.2577
	#observation	612	265	877
Total	Women's education level	13.03347	12.17925	12.72417
	#observation	1494	848	2342

Panel D

College\Attitude to the marriage		0	1	Total
0	Women's education level	11.33139	10.86024	11.13076
	#observation	685	508	1193
1	Women's education level	14.47466	14.15	14.37859
	#observation	809	340	1149
Total	Women's education level	13.03347	12.17925	12.72417
	#observation	1494	848	2342

APPENDIX B



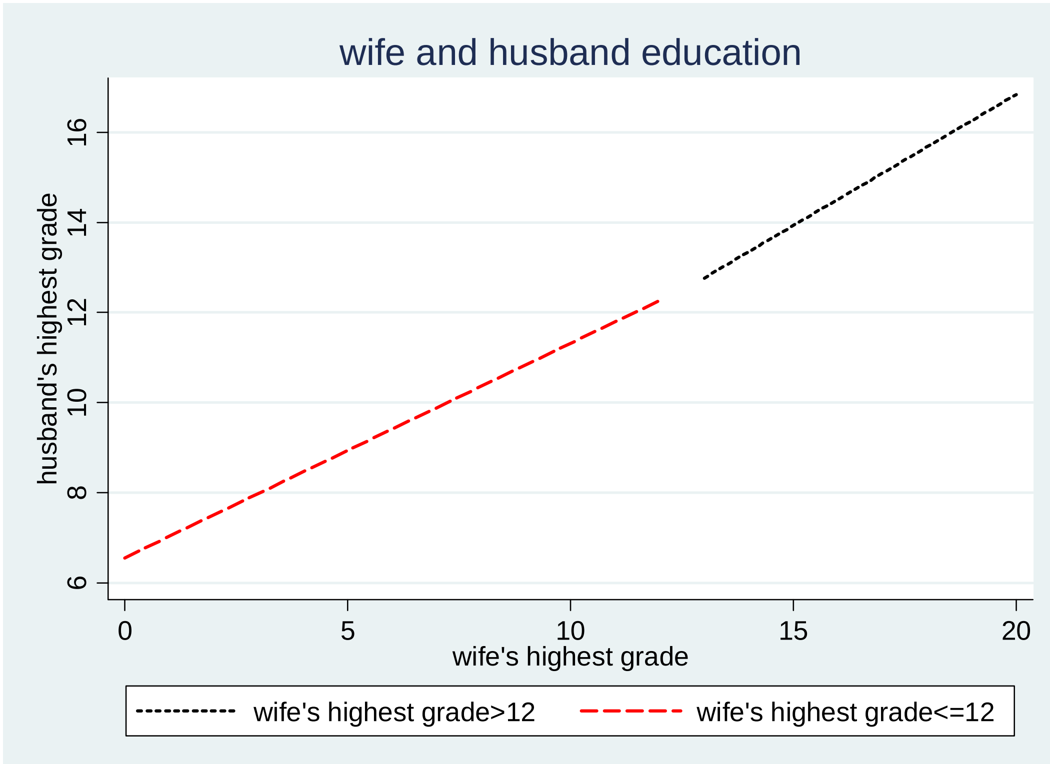
OLS results: Dependent variable: individuals' schooling years

Variable	Non-movers	Movers
Constant	4.0547***	4.3448***
	(0.1333)	(0.3197)
% high school graduates	0.0210***	0.0297***
(race-specific)	(0.0027)	(0.0064)
Observations	3,643	926
R^2	0.016	0.023

Standard errors in parentheses

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

Figure 1: The relationship between the individual schooling years and the county-level race-specific percentages of highs school graduates



Variables	Wife's highest grade>12	Wife's highest grade<=12
Wife's highest grade	0.582***	0.476***
	(0.0282)	(0.0288)
Constant	5.198***	6.553***
	(0.430)	(0.332)
Observations	1995	2459
R-squared	0.176	0.100

Standard errors in parentheses

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

Figure 2: The relationship between the wife's education and husband's education

APPENDIX C

Table C1 Data statistics before and after imputation comparison

	Before		After	
Variables	**Obs**	**Mean**	**Obs**	**Mean**
Mom highest grade	5128	10.8181	5451	10.8120
		(3.0866)		(2.9985)
Dad highest grade	4631	10.9581	5448	10.9032
		(3.8101)		(3.5286)
The oldest sibling's highest grade	3945	11.8063	5446	11.8251
		(3.1702)		(2.6991)
Family size	5454	5.1459	5454	5.1459
		(1.9937)		(1.9937)
Family income	4498	16095.06	5444	16107.60
		(12371.00)		(11326.28)
Safe feeling about the school (1 if feel safe)	4982	0.6200	5451	0.5667
		(0.4854)		(0.4956)
Crime rate	5444	5419.8330	5444	5419.8330
		(3786.1470)		(3786.1470)
Mortality rate	4979	107.0678	5383	109.8763
		(103.1462)		(103.0561)
employment in the education institution	5408	7.8425	5408	7.8425
		(2.5455)		(2.5455)
per capita income	5452	4396.2320	5452	4396.2320
		(903.9928)		(903.9928)
%black	5452	13.4507	5452	13.4507
		(13.9422)		(13.9422)
civil unemployment rate	5452	4.5879	5452	4.5879
		(1.7735)		(1.7735)
%urban	5452	71.9662	5452	71.9662
		(28.2481)		(28.2481)

Table C2-1 Segregation indexes definition

Variable	Dimension of Segregation	Explanation
Absolute Centralization Index (ACE)	measures of centralization	Examines only the distribution of the minority group around the center and varies between -1.0 and 1.0. Positive values indicate a tendency for minority group members to reside close to the city center, while negative values indicate a tendency to live in outlying areas. A score of 0 means that a group has a uniform distribution throughout the metropolitan area
Relative Centralization Index (RCE)	measures of centralization	The relative share of the minority population that would have to change their area of residence to match the centralization of the majority. The index varies between -1.0 and 1.0 with positive values indicating that minority members are located closer to the center than majority, and negative values the reverse. An index of 0.0 indicates that the two groups have the same spatial distribution around the center
Absolute Clustering Index (ACL)	measures of clustering	Expresses the average number of minority members in nearby area units as a proportion of the total population in those nearby area units, where distances between area units are measured from their centroids. It varies from 0.0 to 1.0
Atkinson Index with b=#1 (A1)	measurement of inequality	The Atkinson index allows the researcher to differentially weight area units at different points along the Lorenz curve. The shape parameters 0.1 means minorities are "underrepresented" and this area units will contribute more to segregation index
Entropy Index (H)	measurement of evenness	Measures the weighted average deviation of each area unit from the metropolitan area's "entropy" or racial and ethnic diversity. It varies between 0.0 (when all areas have the same composition as the entire metropolitan area) and 1.0 (when all areas contain one group only).
Gini Index (G)	measurement of evenness	Derived from the Lorenz curve, and varies between 0.0 and 1.0, with 1.0 indicating maximum segregation
Relative Concentration Index (RCO)	measurement of similarly	Takes account of the distribution of the majority group as well. This measure varies from -1.0 to 1.0. A score of 0.0 means that the minority and majority groups are equally concentrated. An index of -1.0 means that the concentration of the majority exceeds that of the minority to the maximum extent, and an index of 1.0 the reverse.

Source: http://www.census.gov/hhes/www/housing/housing_patterns/app_b.html

Table C2-2 Segregation indexes statistics

Variables	Statistics
Absolute Centralization Index (ACE)	0.2549
	(0.3692)
Absolute Clustering Index (ACL)	0.1263
	(0.2013)
Atkinson Index with b=0.1 (A1)	0.0728
	(0.1164)
Entropy Index (H)	0.1697
	(0.2573)
Gini Index (G)	0.2771
	(0.4009)
Relative Centralization Index (RCE)	0.1090
	(0.1877)
Relative Concentration Index (RCO)	0.2178
	(0.3394)

APPENDIX D

This appendix studies the small sample properties of the type-consistent model using simulated samples. In the simulation, the parameter values are given as following:

(1) The main behavior equations for the two different types are given by:

$$\begin{aligned} y_i &= 0.3 - 0.5x_{1i} + 0.26\,x_{2i} + 0.34\,x_{2i} * m_i + u_{1i} && \text{if type} = 1 \\ y_i &= 0.5 - 0.12x_{1i} + 0.6\,x_{2i} - 0.24\,x_{2i} * m_i + u_{0i} && \text{if type} = 0 \end{aligned}$$

(2) The dummy variable m_i is governed by

$$m_i = 1\left(-0.1 + 0.01x_{1i} + 0.15x_{2i} + 0.03x_{3i} - 0.032x_{4i} + v_i > 0\right).$$

(3) The y_i equation and the dummy m_i equation have the following relationship:

$$\operatorname{cov}(u_{0i}, v_i) = 0,\ \ \operatorname{cov}(u_{1i}, v_i) = 0.6$$

(4) Type T is determined by: $T_i = 1\,(0.003 + 1.1z_{1i} + \varepsilon_i > 0)$.

The variables used in the simulation are independently drawn with following distribution:

$$v_i \sim N(0,1);\ \ \varepsilon_\mathrm{i} \sim N(0,1);\ \ \mathrm{u}_{0\mathrm{i}} \sim N(0,1);\ \ \mathrm{u}_{1\mathrm{i}} \sim N(0,1);\ \ \omega_\mathrm{i} \sim N(0,(1-0.64^2)),\ \text{and}$$

$$x_1, x_2, x_3, x_4, y_0, y_1, z_1 \sim N(0,1)$$

I are mainly interested in the estimation bias under different sample sizes. In the simulation, *I* consider four different sample sizes: 3,000, 1,000, 500, and 100. For each sample size, *I* simulate a sample according to above setup and estimate both the Heckman model and the type-consistent model. The difference between the estimated coefficients and the true values is the bias. This process is repeated 500 times for each sample size. Table D1 reports the average biases from the 500 processes.

In this table, it is clear that coefficient estimates using the Heckman sample selection models have large biases, regardless if the $m_i = 1$ sub-sample or the $m_i = 0$ sub-sample is used, and regardless of the sample size. This is not surprising since the data is simulated according to the type-consistent model. For the type-consistent model, coefficient estimates are very close to the true values when the sample size is 3,000. When the sample size is 1,000, the largest bias is 38.9% of the true value, occurring at β_{01} (true value is -0.12, bias is .0467). The bias for the parameter γ_{11} is 13.7% (true value is 0.26, bias is 0.0355). The biases for all other parameters are lower than 5% of the true values. Therefore, one has to be cautious when the sample size is around 1,000.

When sample size is 500, the bias is rather large for some coefficients. For example, for β_{01}, the bias is 185.5% of the true value (the true value is -.12, while the bias is .2226). For γ_{11}, the bias is 65.0% of the true value (bias is 0.1691 while the true value is .26). Therefore, using the type-consistent model when sample size is 500 is problematic. When sample size is 100, estimates are clearly no longer reliable.

Table D1-1 Small sample bias of the type-consistent model

Coef	True value	Obs	Heckman Model Sub-sample $m_i = 0$ Bias	Heckman Model Sub-sample $m_i = 1$ Bias	Type-consistent Model Bias
β_{11}	-0.5	100	-0.3890	-0.3747	0.3254
γ_{11}	0.26	100	0.4734		0.6299
$\gamma_{11}+\gamma_{12}$	0.6	100		-0.2616	0.0276
θ	0.6	100	-0.3615	-0.2616	0.0082
β_{1}	0.3	100	-0.4213	1.4392	0.1877
β_{01}	-0.12	100	1.5460	1.6054	1.3851
γ_{01}	0.6	100	-0.3615		0.2625
$\gamma_{01}+\gamma_{22}$	0.36	100		0.2306	0.1120
β_{0}	0.5	100	-0.6528	0.4635	0.0945
β_{11}	-0.5	500	-0.3751	-0.3594	0.0472
γ_{11}	0.26	500	0.5744		0.1691
$\gamma_{11}+\gamma_{12}$	0.6	500		-0.2362	0.0498
θ	0.6	500	-0.3178	-0.2362	0.0021
β_{1}	0.3	500	-0.0752	0.7036	0.0553
β_{01}	-0.12	500	1.6039	1.6691	0.2226
γ_{01}	0.6	500	-0.3178		0.0588
$\gamma_{01}+\gamma_{22}$	0.36	500		0.2731	0.0636
β_{0}	0.5	500	-0.4451	0.0222	0.0376

Table D1-2 Small sample bias of the type-consistent model

			Heckman Model		Type-consistent
	True		Sub-sample $m_i = 0$	Sub-sample $m_i = 1$	Model
Coef	**value**	**Obs**	**Bias**	**Bias**	**Bias**
β_{11}	-0.5	1000	-0.3834	-0.3894	0.0164
γ_{11}	0.26	1000	0.5996		0.0355
$\gamma_{11}+\gamma_{12}$	0.6	1000		-0.2175	0.0032
θ	0.6	1000	-0.3069	-0.2175	0.0045
β_1	0.3	1000	-0.1095	0.7190	0.0086
β_{01}	-0.12	1000	1.5692	1.5441	0.0467
γ_{01}	0.6	1000	-0.3069		0.0087
$\gamma_{01}+\gamma_{22}$	0.36	1000		0.3042	0.0150
β_0	0.5	1000	-0.4657	0.0314	0.0098
β_{11}	-0.5	3000	-0.3821	-0.3761	0.0033
γ_{11}	0.26	3000	0.6289		0.0015
$\gamma_{11}+\gamma_{12}$	0.6	3000		-0.2015	0.0010
θ	0.6	3000	-0.2941	-0.2015	0.0000
β_1	0.3	3000	0.1251	0.3489	0.0064
β_{01}	-0.12	3000	1.5748	1.5996	0.0059
γ_{01}	0.6	3000	-0.2941		0.0052
$\gamma_{01}+\gamma_{22}$	0.36	3000		0.3308	0.0062
β_0	0.5	3000	-0.3250	-0.1907	0.0053

APPENDIX E

To derive the likelihood function, it is necessary to have joint density. However, before considering the joint density of College and *Marriage*, let's look at the conditional density first:

$$
\begin{aligned}
&\Pr\left(Marriage=1 \mid College=0\right)\\
&=\Pr\left(Marriage=1 \mid T=0, College=0\right)\Pr(T=0)\\
&\quad+\Pr\left(Marriage=1 \mid T=1, College=0\right)\Pr(T=1) \qquad\qquad \text{(C-1)}\\
&=\Pr\left(X\beta_{01}+\beta_{02}educ+\zeta_0>0 \mid College=0\right)(1-\Phi(w\pi))\\
&\quad+\Pr\left(X\beta_{11}+\beta_{12}educ+\zeta_1>0 \mid College=0\right)\Phi(w\pi)
\end{aligned}
$$

For the type 0, λ_0 is uncorrelated with ξ_0. Therefore, the first part of (A-1) is:

$$
\begin{aligned}
&\Pr\left(Marriage=1 \mid T=0, College=0\right)\\
&=\Pr\left(X\beta_{01}+\beta_{02}educ+\zeta_0>0 \mid T=0, College=0\right)\\
&=\Pr\left(\zeta_0>-X\beta_{01}-\beta_{02}educ\right)\\
&=\Phi\left(X\beta_{01}+\beta_{02}educ\right)
\end{aligned}
$$

The second part of (A-1) is more complicated, because λ_1 is correlated with ξ_1 in this case:

$$
\begin{aligned}
&\Pr\left(Marriage=1 \mid T=1, College=0\right)\\
&=\Pr\left(X\beta_{11}+\beta_{12}educ+\zeta_1>0 \mid College=0\right)\\
&=\Pr\left(X\beta_{11}+\beta_{12}educ+\rho\lambda+v>0 \mid College^{*}\le 0\right)\\
&=E\left[\Phi\left(\frac{X\beta_{11}+\beta_{12}educ+\rho\lambda}{\sqrt{1-\rho^2}}\right)\middle|\,\lambda<-\theta_1 z\right]\\
&=\frac{1}{\Pr\left(\lambda<-\theta_1 z\right)}E\left[\Phi\left(\frac{X\beta_{11}+\beta_{12}educ+\rho\lambda}{\sqrt{1-\rho^2}}\right),\lambda<-\theta_1 z\right]\\
&=\frac{1}{\Phi\left(\theta_1 z\right)}\int_{-\infty}^{-\theta_1 z}\Phi\left(\frac{X\beta_{11}+\beta_{12}educ+\rho\lambda}{\sqrt{1-\rho^2}}\right)\phi\left(\lambda\right)d\lambda
\end{aligned}
$$

Based on the conditional density, I can calculate the joint density. For the case *Marriage* =1 and *College* =0, the joint density is

Case I: *Marriage* =1 and *College* =0

$$\Pr\left(Marriage=1, College=0\right)=\Pr\left(Marriage=1 \mid College=0\right)\Pr\left(College=0\right)$$
$$=\begin{bmatrix}\Pr(Marriage=1 \mid T=0, College=0)\Pr(T=0)\\ +\Pr(Marriage=1 \mid T=1, College=0)\Pr(T=1)\end{bmatrix}\Pr\left(College=0\right)$$
$$=\begin{bmatrix}\Phi(w\pi)\Pr(Marriage=1 \mid T=0, College=0)\\ +(1-\Phi(w\pi))\Pr(Marriage=1 \mid T=1, College=0)\end{bmatrix}\Pr\left(College=0\right)$$
$$=(1-\Phi(w\pi))\Phi\left(X\beta_{01}+\beta_{02}educ\right)(1-\Phi\left(\theta_0 z\right))$$
$$+\Phi(w\pi)\int_{-\infty}^{-\theta_1 z}\Phi\left(\frac{X\beta_{11}+\beta_{12}educ+\rho\lambda}{\sqrt{1-\rho^2}}\right)\phi\left(\lambda\right)d\lambda$$

By using similar method, the joint density for the rest of three cases can be written as:

Case II: *Marriage* =0 and *College* =0

$$\Pr\left(Marriage=0, College=0\right)=\Pr\left(Marriage=0 \mid College=0\right)\Pr\left(College=0\right)$$
$$=\begin{bmatrix}\Pr(Marriage=0 \mid T=0, College=0)\Pr(T=0)\\ +\Pr(Marriage=0 \mid T=1, College=0)\Pr(T=1)\end{bmatrix}\Pr\left(College=0\right)$$
$$=\begin{bmatrix}\Phi(w\pi)\Pr(Marriage=0 \mid T=0, College=0)\\ +(1-\Phi(w\pi))\Pr(Marriage=0 \mid T=1, College=0)\end{bmatrix}\Pr\left(College=0\right)$$
$$=\left(1-\Phi(w\pi)\right)\Phi\left(-X\beta_{01}-\beta_{02}educ\right)(1-\Phi\left(\theta_0 z\right))$$
$$+\Phi(w\pi)\int_{-\infty}^{-\theta_1 z}\Phi\left(\frac{-X\beta_{11}-\beta_{12}educ-\rho\lambda}{\sqrt{1-\rho^2}}\right)\phi\left(\lambda\right)d\lambda$$

Case III : *Marriage* =0 and *College* =1

$$
\begin{aligned}
&\Pr\left(Marriage=0, College=1\right)=\Pr\left(Marriage=0 \mid College=1\right)\Pr\left(College=1\right)\\
&=\begin{bmatrix}\Pr(Marriage=0 \mid T=0, College=1)\Pr(T=0)\\ +\Pr(Marriage=0 \mid T=1, College=1)\Pr(T=1)\end{bmatrix}\Pr\left(College=1\right)\\
&=\begin{bmatrix}\Phi(w\pi)\Pr(Marriage=0 \mid T=0, College=1)\\ +(1-\Phi(w\pi))\Pr(Marriage=0 \mid T=1, College=1)\end{bmatrix}\Pr\left(College=1\right)\\
&=\left(1-\Phi(w\pi)\right)\Phi\left(-X\beta_{01}-\beta_{02}educ\right)\Phi\left(\theta_0 z\right)\\
&\quad+\Phi(w\pi)\int_{-\theta_1 z}^{\infty}\Phi\left(\frac{-X\beta_{11}-\beta_{12}educ-\rho\lambda}{\sqrt{1-\rho^2}}\right)\phi\left(\lambda\right)d\lambda
\end{aligned}
$$

Case IV : *Marriage* =1 and *College* =1

$$
\begin{aligned}
&\Pr\left(Marriage=1, College=1\right)=\Pr\left(Marriage=1 \mid College=1\right)\Pr\left(College=1\right)\\
&=\begin{bmatrix}\Pr(Marriage=1 \mid T=0, College=1)\Pr(T=0)\\ +\Pr(Marriage=1 \mid T=1, College=1)\Pr(T=1)\end{bmatrix}\Pr\left(College=1\right)\\
&=\begin{bmatrix}\Phi(w\pi)\Pr(Marriage=1 \mid T=0, College=1)\\ +(1-\Phi(w\pi))\Pr(Marriage=1 \mid T=1, College=1)\end{bmatrix}\Pr\left(College=1\right)\\
&=\left(1-\Phi(w\pi)\right)\Phi\left(X\beta_{01}+\beta_{02}educ\right)\Phi\left(\theta_0 z\right)\\
&\quad+\Phi(w\pi)\int_{-\theta_1 z}^{\infty}\Phi\left(\frac{X\beta_{11}+\beta_{12}educ+\rho\lambda}{\sqrt{1-\rho^2}}\right)\phi\left(\lambda\right)d\lambda
\end{aligned}
$$

After deriving all of joint densities, the likelihood function is:

$$
\begin{aligned}
l=&\sum_{Marriage=1, College=0}\log\left(\Pr\left(Marriage=1, College=0\right)\right)\\
&+\sum_{Marriage=0, College=0}\log\left(\Pr\left(Marriage=0, College=0\right)\right)\\
&+\sum_{Marriage=1, College=1}\log\left(\Pr\left(Marriage=1, College=1\right)\right)\\
&+\sum_{Marriage=0, College=1}\log\left(\Pr\left(Marriage=0, College=1\right)\right)
\end{aligned}
$$

VITA

Name: Yingning Wang

Address: 3009 Allen, Department of Economics, TAMU, College Station, TX 77843-4228

Email Address: wangyingning@tamu.edu

Education: B.A., International Finance, Central South University, 1998

M.A., International Finance, Fudan University, 2001

Ph.D., Economics, Texas A&M University, 2009