PHD

Private returns to education: earnings, health and well-being

Qiu, Tian

Award date:
2007

Awarding institution:
University of Bath

Link to publication

Private Returns to Education: Earnings, Health and Well-being

Tian Qiu

A Thesis Submitted for the Degree of Doctor of Philosophy

University of Bath

Department of Economics and International Development

September 2007

Signature: ____

Contents

List of Tables

List of Figures

Abbreviations

BMI	Body Mass Index
CHIP	Chinese Household Income Project
CHNS	China Health and Nutrition Survey
CLMRP	Chinese Labour Market Research Project
CMS	Cooperative Medical System
COEs	collectively-owned enterprises
FDI	foreign direct investment
FE	fixed effects
GDP	Gross Domestic Product
GIS	Government Insurance Scheme
GLS	Generalised Least Squares
GSOEP	German Socio-Economic Panel
GSS	General Social Survey
IQ	intelligence quotient
IRR	internal rate of return
IV	instrumental variable
LIS	Labour Insurance Scheme
LM	Lagrange multiplier test
MES	modern enterprise system
NPV	net present value
PSID	Panel Study of Income Dynamics
RE	random effects
SES	socio-economic status
SEZs	special economic zones
SOEs	state-owned enterprises
SRH	self-reported health
SSH	strong screening hypothesis
SWB	subjective well-being

WHO	World Health Organization
WSH	weak screening hypothesis
WTO	World Trade Origination

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor John Hudson, for his extraordinary support and help in the past four and a half years. Without his illuminating instructions and intellectual comments, this thesis would not have been possible.

I would also like to express my gratefulness to Dr. Pamela Lenton (external examiner) and Professor John Sessions (internal examiner) for their detailed corrections and useful comments.

My personal appreciation goes to Tao who has given me strength and carried me in the last year of my study. Also I appreciate the faculty members in the Economics department and my colleagues who contributed disputable viewpoints in research seminars and have given me many constructive suggestions.

Finally, and most importantly, I would like to take this opportunity to express my deepest thanks to my dear parents for their endless love and encouragement through my entire life.

Abstract

This thesis is focused on economic returns to education in China. It takes education as a key point to develop three aspects: earnings returns to education, the influence of education as one of the social-economic factors associated with the Body Mass Index (BMI) related to health, and the determinants of subjective well-being. A panel data-based analysis including continuous and discrete dependent variables (ordered probit/probit) is used in this research.

The first part examines the earnings returns to education in urban China for four years covering the period 1989 to 2000. We find, in common, with others that such returns were small in 1989, but have increased steadily since then. We also find that the returns for women exceed those for men and go some way to reducing the gender earnings gap. Crucially, however the returns to education decline with the length of time since the individual left school which is consistent with the hypothesis that education enhances ability and skills which in turn enhances earnings, but that the value of such skills deteriorates over time. Finally we find evidence for gravity effects by which earnings decline as distance from Beijing, and more noticeable, Shanghai increases.

The aim of the second part is to examine the impact of socio-economic status (SES) on the BMI, a formula based on the ratio of height to weight, linked to health, using a four-year (1991, 1993, 1997 and 2000) panel data set. To an extent we confirm the results with respect to the linkage between SES and health found for other countries. However,

instead of using the existing specification of BMI, we explore the healthy BMI range based on a self-reported measure of health in China. This leads to a slightly different formulation for the BMI and a substantially different healthy range. We also find that this healthy BMI has a significant impact on health together with SES. Because of potential simultaneity between education and health we estimate a relationship between SES and health change. We find a significant relationship between education and changes in health status.

The final part studies both happiness and life satisfaction in mainland China. We explore the extent to which SES and social capital influence subjective well-being. The results for happiness and life satisfaction are similar, but not identical. To an extent we confirm the results of others with respect to other countries. Hence we find a U-shaped relationship with respect to age and positive influences of income and health on well-being. We also include a variable which reflects the degree of choice/control over their lives people feel they have. Crucially, for social capital variables, we find that individuals who are involved in more voluntary organizations have higher levels of happiness, and those who are a member of Communist party are also more satisfied with their lives. Finally, education has a limited positive impact on subjective well-being, however, it is also the most significant determinant of social capital variables across individuals.

In the thesis we specifically discuss the problem of endogeneity which is traditionally tackled by the use of some instrumental variable method. Recently much of the work in this genre including work relating to education, has been criticised from the perspective of weak instruments. Throughout we suggest alternative approaches and each is specific to the context in which it is used. Each of these alternatives is in itself based on certain assumptions which can in turn be questioned. Their value lies more in that they present extra evidence on the impact of education, rather than they unambiguously provide a solution to the endogeneity problem. Economics is not an exact science and it is the accumulation of evidence which is important. In our case our evidence is that education matters, and it matters not just with respect to earnings, and by implication productivity, but also with respect to health and subjective well-being.

Chapter 1 Introduction

The importance of human capital as a source of economic growth has been recognized in economic literature dating back to Adam Smith (1776). However, it was not until the 1960s that the human capital theory was formally developed by Schultz (1961), Becker (1964) and Mincer (1974). These economists rekindled the human capital concept by affirming its links with economic growth and by emphasizing its importance in explaining earnings differentials (Laroche, *et al.*, 1999). Thus the human capital contributions focused largely on differential earnings of individuals as they related to skills. The most natural way to identify differential skills was the amount of education completed by individuals (Hanushek and Welch, 2006). A benchmark model for the earning returns to education has been developed by Mincer (1974), and a great deal of work based on this model has been done for many countries.

For decades, the primary argument for justifying education has been based on its direct impact on labour market productivity and earnings. Yet it is widely perceived that the effects of education spread beyond this to include social benefits for both individuals and societies, such as better health, higher levels of welfare or lower crime rates, etc. Such benefits result in better equipped individuals to take care of themselves and consequently create a better society in which to live. According to Behrman, *et al*. (1997), the quantification of these social benefits is difficult, but more systematically analyzing these benefits not only would improve our understanding of the full effects of education but also would improve the informal basis for considering policies related to education.

This is the central focus of our thesis with specific reference to China focusing on three specific issues. In the next section, we generally describe the background and theoretical issues related to these three topics. In section 2, we outline more specifically the purpose and the key questions of the thesis. We also briefly introduce the data sets we use in our empirical work. The last section outlines the following chapters.

1.1. Background and Theories

1.1.1. Human Capital

The concept of human capital is central to much of the research on the economic returns of education, we thus start with briefly introducing the human capital theory. It is the theoretical foundation of our first empirical topic. In human capital theory, education is an investment of current resources in exchange for future returns. The benchmark model for the development of empirical estimation of the returns to education is the key relationship derived by Mincer (1974). Empirically, this involves regressing the natural logarithm of earnings against educational attainment and work experience. Note that experience is included as a quadratic term to capture the concavity of the earnings profile. Mincer's derivation of the empirical model implies that, under the assumptions made, the coefficient of education can be considered as the private monetary return to education as well as being the proportionate effect on wages of an increment to education (Harmon, *et* *al.*, 2003).

However, there exist difficulties in this analysis. For example, the coefficient on years of education may not (only) reflect the effect of education on productivity if it is correlated with unobserved characteristics that are also correlated with wages. In this case, the education coefficient would reflect both the effect of education on productivity and wages and the effect of the unobserved variable that is correlated with education. For example, 'ability' may be unobservable and may be correlated with the ability to earn money in the labour market.

There have been a number of approaches dealing with this problem. Firstly, measures of ability have been incorporated to proxy for unobserved effects, such as intelligence quotient (IQ) scores. The inclusion of direct measures of proxing ability should reduce the estimated education coefficient, so that the coefficient on education then captures the effect of education alone. The usual finding is that the coefficient on education is not strongly affected by the inclusion of additional right-hand-side variables (see, e.g., Weiss, 1995). Secondly one might explore within-twins or within-siblings differences in wages and education with the assumption that unobserved effects are additive and common within twins. This approach is clearly restricted by data availability. A final, and the most common, approach deals directly with the education-earnings relationship in a two-equation system by exploiting instrumental variables (IV) that affect education directly, but earnings indirectly only through the impact on education. However, the IV method can not always be applied to the empirical work as either necessary instruments may not be available, or they are subject to weak instruments bias.

1.1.2. Health Capital

Building on the human capital theory, Grossman (1972) provides a formal model to analyze health capital. According to this approach, the main distinction between health and education is that health increases income through adding healthy working days while education works through improving productivity (Zhao, 2005). Grossman's model has become the standard model to study health demand and health determinants, and our second topic identifies such determinants in China applying Grossman's model.

An extensive review of the literature conducted by Grossman and Kaestner (1997) suggests that years of formal education completed is the most important correlate of good health. This finding emerges whether health levels are measured by self-evaluation of health status, mortality rates or physiological indicators of health. In a broad sense, the observed correlation between education and health may be explained in one of three ways according to Grossman (2000). Firstly, there is a causal relationship that runs from increases in education to increases in health; secondly, the direction of causality runs from better health to more education; and thirdly, there is no causal linkage between education and health, instead, differences in one or more 'third variables' affect both education and health in the same direction. Similar to the last section, the 'third variable' explanation is relevant if the unaccounted factors which affect health are correlated with education, such as inherent health. It is clear that the second and third explanation challenged the causal relationship between education and health.

Another important aspect which is commonly involved in this context is the concept of a healthy weight range, which is based on a measurement known as the Body Mass Index (BMI). The BMI is calculated as weight in kilograms over height in metres squared (weight (kg)/height $(m)^2$). The World Health Organization (WHO) has devised a classification where by persons with BMIs below the range 19-25 are considered underweight, those with BMIs above this range are considered overweight or “at risk”, and those with BMIs greater than or equal to 30 are considered obese (Bell, *et al.*, 2002). The general finding suggested in the literature is that a healthy BMI tends to lead to better health (see, e.g., Duetz, *et al.*, 2000; Beydoun and Popkin, 2005). However, there is a possibility that different ethnic groups have similar health risks at different levels of BMI. For instance, instead of more than (≧) 25, the overweight status among Chinese is defined as a BMI≧24 (Wildman, *et al.*, 2004). The issues relevant to BMI are also related to our second topic. In particular we seek to analyse i) the validity of the basic formula, ii) whether it is applicable to all people regardless of age and gender, iii) the parameters of the healthy range and iv) the socio-economic determinants of the BMI.

1.1.3. Social Capital

Helliwell (2001) argued that the working definition of social capital refers to the networks, norms and understandings that facilitate cooperative activities within and among groups of individuals. Similarly, Halpern (2005) argued that most forms of social capital consist of a *network*, a cluster of *norms, values, and expectancies* that are shared by group members; and *sanctions* that help to maintain the norms and network. Much of the interest in social capital amongst economists has been fuelled by a definition that includes not only the structure of networks and social relations, but behavioural dispositions (such as trust, reciprocity, and honesty) and institutional quality measures (Woolcock, 2001). Recent work has sought to break the notion of social capital into different sub-types, and perhaps the most important of these distinctions is between 'bonding' and 'bridging' (Gittell and Vidal, 1998). As Putman (2000) argued, some forms of capital are (i.e., *bonding* social capital), by choice or necessity, inward looking and tend to reinforce exclusive identities and homogeneous groups; other networks (i.e., *bridging* social capital) are outward looking and encompass people across diverse social cleavages. However, many actual situations are characterized by social capital that mixes these two types in different proportions (Helliwell, 2001).

The interest in social capital has been driven in recent academic research, although it is difficult to ideally measure it as is e.g. years of education for human capital. A large number of studies have emerged documenting a relationship between the form and quality of people's networks and a range of important outcomes, including economic growth, health and welfare. The standard measures of social capital used in these studies are mainly the number of non-professional organizations that an individual belongs to and individuals' general trust levels, namely the extent to which most people in a given community, region or nation trust each other (Glaeser, 2001). Furthermore, in all studies of differences amongst individuals as to the extent of their connectedness and trust, education levels are uniformly the strongest and most consistent of explanatory factors.

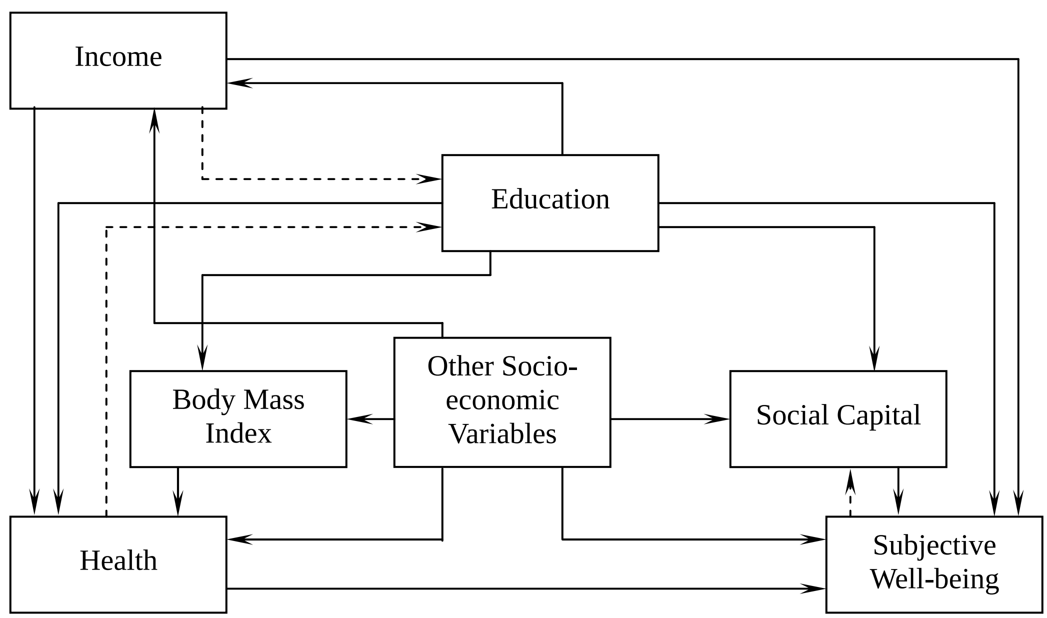
According to Putman (2001), although social capital has a multitude of potentially measurable consequences, we are not yet in a position to rule out all other explanations for these patterns. For example, it is also possible that the causality runs to social capital instead of from social capital. This is an important and indeed interesting enough field that it is worth paying more attention to. The linkage between some social capital measures and subjective well-being is closely related to our third topic.

1.2. Purpose and Approach of the Thesis

The main areas of concern in this thesis lies with the impact of education on different outcomes in China based on the individual level. They correspond to the three main parts to the thesis, i.e. i) earnings returns to education, ii) the influence of socio-economic status (SES) and BMI on self-reported health, and iii) the impact of SES and social capital on subjective well-being. Figure 1-1 summarizes the thesis.

A central question in this thesis is how causal inferences can be drawn when only limited observational data are available. Motivation for causal inference in econometric models is especially strong when the causal variables themselves reflect individuals' decisions and hence are potentially endogenous. Therefore the inferences on such causal relationships need to be drawn based on both economic theory and econometric techniques. The problem of endogeneity is traditionally tackled by the use of some instrumental variable method, although recently much of the work in this genre has been criticised from the perspective of weak instruments. Throughout the thesis, we suggest alternative approaches, and their value lies more in that they present extra evidence on the impact of education, rather than they unambiguously provide a solution to the endogeneity problem. Indeed it is reasonable to argue that, such is the nature of the endogeneity problem with respect to education based on a lack of data on e.g. pure ability, there is no ideal solution.

Figure 1-1 Diagram of the Thesis Representation



Notes: The broken lines indicate the possibility of reverse causality or the 'third variables' affecting both outcomes.

Apart from education a second focus of the thesis is on China. How does education impact on earnings, health and well-being in China? Are these impacts comparable with those found in other studies, primarily of developed market economies? This is important if we are to understand how the results of such research can be generalized across people in different countries. But in addition China is emerging as a major economic power, and an understanding of that country and its people is also important.

The approach in this thesis is both structural and parametric. Economic theory is used to formulate the models and to provide empirically testable hypotheses in some cases. The emphasis is on the empirical analysis. A lot of effort has been devoted to stretching the statistical methods to obtain consistent estimates.

The data we use for the first two parts, corresponding to the fourth and fifth chapters respectively of the thesis, comes from the Household Survey and Physical Examination in China Health and Nutrition Survey (CHNS). The CHNS was collected mainly by the Carolina Population Centre and it is designed to examine the effects of the health, nutrition, and family planning policies and programs implemented by national and local governments and to see how the social and economic transformation of Chinese society is affecting the health and nutritional status of its population. The sample from the Household Survey for the first part of the work covers the years of 1989, 1993, 1997 and 2000, and the Physical Examination for the second part covers the years of 1991, 1993, 1997 and 2000. Both of these are panel data sets and they cover much of the period of China's economic reform process, thus allowing for a comparison of some major variables across the nineties in China.

The third part utilizes cross-sectional microdata of World Values Survey (1990, 1995-1997, 1999-2000), which is designed to enable a cross-national, cross-cultural comparison of values and norms on a wide variety of topics. For the relevance of this thesis, these data contain information on attitudes to quality of life, social involvements and trust levels which were not commonly available in a Chinese context.

1.3. Chapter Outline

The next chapter reviews China's economic reforms and existing regional inequality related to income, health care and education. China's transition from a centrally-planned to a market economy began at the end of 1978 motivated by the realizations of the need for a rapid expansion of an economy characterized by low allocative and technical efficiency. Economic reforms have brought significant changes to China's labour market and market determined wage rates. However, it has also led to a growing inequality of income, specifically between urban and rural areas, as well as between inland and coastal provinces. This chapter presents the transitional background to evaluate the implications of our econometric results for the returns to education in China.

The theoretical framework and empirical background relevant to educational returns are presented in chapter 3. We mainly focus on two representative theories: human capital theory and the concept of health capital. Specifically we examine the possibilities that education may not have a causal impact on earnings and health. In addition, we also generally discuss the measurement and measurable consequences of social capital. The discussion of these theoretical and empirical issues introduces the main research themes and corresponding hypotheses, which are analyzed and tested in the following three chapters.

Chapter 4, 5 and 6 form the main part of the thesis. The empirical findings from the analysis of economic returns to education related to earnings, health and subjective well-being are presented in these three chapters. Chapter 4 estimates the private returns to investment in years of education in urban China. Within the commonly used logarithmic wage functional framework, we focus on the potential biases, mainly unmeasured ability, in the endogeneity of educational choice which is mostly ignored for developing countries. We deal with this in a slightly different way to others as we do not have suitably strong instrumental variables available in our dataset. We assume that the impact of education itself on earnings may be because education either (i) *enhances* ability and productivity or (ii) because it *reflects* ability. If the former is true then in a changing world, skills, including educationally enhanced skills, will date and hence the value of education in enhancing earnings should decline as the period since formal education completed increases. We found strong evidence for this in the regression results with the assumption that the value of education as a signal does not decline with age.

Chapter 5 examines the impact of SES and BMI on self-reported health. Specifically, we attempt to establish the causal nature of the relationship between education, income and health. Based on the possibility of unobserved individual heterogeneity, we apply a recently-developed statistical method for estimating an ordered logit model with fixed-effects as well as an ordered probit model with random-effects. We find, specifically for the income, the magnitude and significance of the coefficients in random-effects and fixed-effects models are very close. In addition, also as a robustness check, using an approach based on the relationship between SES and the change of health status we find the SES characteristics of people whose health status is deteriorating.

BMI is more closely related to health behavior than self-reported health, and it is another important aspect included in this chapter. We question the general calculation of BMI (weight(kg)/height(m)^2), and we extend it to include the impact of both gender and age. Based on this analysis we calculate alternative BMIs. In addition, we also examine the 'healthy range' in China based on our new measures of BMI and further estimate its impact on self-reported health together with SES.

Education and health are the two most important investments individuals make. Both investments make individuals more productive, and they also have a considerable impact on their subjective level of well-being. Chapter 6 estimates the impact of SES and social capital on happiness and life satisfaction for China. As throughout the thesis, we also focus on the impact of education on social capital. We use social involvements and the trust level to measure the social capital. Specifically, we separate participation in voluntary associations and political parties based on the Chinese context. This chapter thus adds one more piece to the accumulating international evidence on subjective well-being.

In the concluding chapter, we sum up all the findings and explore the policy implications using the evidence revealed in this thesis for educational practitioners and policy makers.

Chapter 2 Economic Reforms in China

2.1. Introduction

The realizations of the potential for rapid expansions of an economy with low allocative and technical efficiency motivated the economic reforms inaugurated in 1978. These reforms were the stimulus for economic development in China, which firstly began in rural areas. The key to the reforms in rural areas was the dissolution of communes across the country and the implementation of the household responsibility system, which provided new opportunities for farm families to take control of their labour allocation decisions. For example, rural households were allowed to accumulate funds to invest in small-scale industry. Non-farm employment was largely increased and a massive flow of migrant workers from the rural areas to the urban areas began in the mid 1980s (Meng, 2000). Since 1984, the focus of the reforms has gradually shifted to urban areas. The reform of state-owned enterprises (SOEs) is seen as the core of urban reform. During the 1990s, Deng introduced the concept of the ‘socialist market economy with Chinese characteristics’ by transforming the ownership system into a mixed structure with the public sector dominant and various types of ownership coexistent (Guo, 2003). In 1990, the share of total urban employment in state units was 62.3%, but by 2000 it had declined to 38.1% [1]. Meanwhile, beginning in 1997, the government moved forward with

[1] Data come from China Statistical Yearbook, collected by the National Bureau of Statistics of China, 2001.

aggressive restructuring and privatization of SOEs, leading to substantial layoffs, retirements, and exits from the labour force (Maurer-Fazio, 1999).

Amongst the developing countries, China has made rapid progress in economic growth, and its annual GDP growth rate has increased on average by 9.5% each year between 1989 and 2005. At the same time, economic reforms have had a great impact on income distribution which led to the growing inequality of income in the 1990s (Uchimura, 2005). There are large inequalities in wage income between urban and rural residents, as well as between different regions. As Han (1998) argued, the urban-rural inequality has been the biggest contributor to the whole problem of inequality in China followed by inter-regional inequality. A key dimension of inter-regional inequality is that between inland and coastal provinces. In addition, there are other dimensions of regional inequality on human development related to income inequality, such as health care and education attainment. The reasons for this are related to both economic and political factors.

This chapter will proceed as follows: In section 2, we generally review the situation of China's labour market in the pre-reform era. Section 3 focuses on the two primary aspects of China's reforms pertaining to the labour market and to the wage-setting system and ownership system for SOEs. In section 4, we investigate the main causes of inter-provincial inequality and rural-urban inequality based on income, health care and education attainment. We conclude this chapter in section 5.

2.2. Labour Market in Pre-reform China

2.2.1. Pre-reform Labour Arrangements and Employment

Before economic reform was initiated in 1978, China did not have labour markets in the conventional sense or as assumed in human capital theory. In the late 1950s to the late 1970s, under the planned allocation of labour, major employers such as SOEs, collectively-owned enterprises (COEs) or government department were not allowed to recruit freely. The labour allocation process began with the central and local government's annual plan for labour quotas. The distribution between employers and employees was totally done through the government labour bureaus without much reference to the needs of individual enterprises or the characteristics of workers. Individuals were not allowed to find jobs by themselves, and the wages and salaries of workers and managers were determined not by their performance but by their education, age, position, and other criteria according to a national wage scale (Lin, *et al*., 1996). In other words, labour at that time was not considered a commodity and wages were not perceived as being the price of labour (Meng, 2000). Furthermore, the state enterprises remitted all their profits to the state and the state budget also would cover all losses incurred by enterprises. The employers had no autonomy in the employment/dismissal of employees as full employment and lifetime tenure, at least in SOEs, were typical features of this system. Thus a person's first job was normally his/her only job, the so called 'iron rice bowl[2]'. According to Meng (2000), by having such an authoritarian structure of job

---

[2] According to Ding and Warner (2001), the 'iron rice bowl' was characterized by a system of

allocation, and by restricting rural-urban migration, the ideological goal of full employment was achieved when the economy was running smoothly[3].

China's population remains predominantly in rural areas and over 80 percent were classified as rural by the 1980 census. Rural-urban labour migration was strictly restricted, as well as inter-regional migration. The restrictions on labour mobility were implemented through the household registration system (*Hukou*) which was set up in the mid-1950s. People were forced to register with local authorities to gain residency, and a place of work, and spend their lives in the place where they were born. In short, individuals were tied to their birthplace. For example, urban enterprises were restricted from recruiting labour from another province unless labour could not be found locally (Brooks and Tao, 2003). More specifically, rural people were excluded from working in the cities, and received no social benefits. Only approved urban registered residents were allowed to live and work in urban areas. More important, only urban *Hukou* holders were entitled to receive welfare benefits including housing, medical care, pensions and sometimes schooling and hospital treatment, provided mostly by the SOEs as an internal social security system (Meng, 2000). It is clear that a great imbalance existed between the treatment of the rural labour force compared to that of the urban one. There were some exceptions, such as rural university graduates or army officials, who could gain urban *Hukou* residency. Before the reforms, these seemed to be the only opportunity for young people to leave the rural registration. This system effectively constrained the development

---

unified job allocation, guaranteed life-long employment, and cradle-to-grave welfare.

[3] The periods, when the economy was running smoothly, did not include the Great Leap Forward (in the late 1950s and early 1960s) and the Cultural Revolution (1966-1969). For detailed discussions see Meng (2000).

of a national labour market. Meng (2000) argued that these restrictions generated serious problems for the Chinese economy including considerable hidden unemployment in the countryside and low productivity and inefficiency in the cities.

2.2.2. The Urban Grade Wage System

As mentioned, China's urban labour market was characterized by a system of centrally fixed wages and lifetime secured employment. The grade wage system which was introduced in 1950s was based on the Soviet model, and continued with hardly any changes until the late 1970s. This complicated wage system can be roughly broken into three sub-systems: i) the eight-scale wage structure for production workers, ii) the post or job-type wage system, for employees mainly in light and service industries; and iii) the responsibility wage system, mainly for administrative and managerial personnel in enterprises[4] (Ding and Warner, 2001). Wage grades were based on personal criteria such as seniority and qualifications, including educational levels and work experience (Shan, 1991). This wage-setting system provided a precise wage for each individual based on the predeterminated wage grades (according to these personal criteria). Enterprises had no right to determine their employees' wages, instead, they were given an annual total wage bill quota by the Ministry of Labour and Personnel, and this wage bill quota was a direct function of enterprises' employment allocation, i.e. the fixed employment quota multiplied by the precise level of wage for each individual employed in the enterprise (Meng, 2000). Further, the wage levels differed by regions. For example, Beijing, the

[4] According to Kidd and Meng (2001), this wage system consists of eight levels for factory workers (ranking in ascending order) and 25-28 levels for carders (ranking in descending order).

capital of China, was ranked in region 6, whereas Shanghai was a more expensive region and was placed in region 8. According to Meng (2000), the higher the rank, the higher the wage levels.

China has adhered to a ‘low-wage policy’ to increase capital accumulation for investment in heavy industries. Wage increases were awarded only according to government regulations apart from accruing to seniority[5]. Employees in industrial enterprises could also get a uniform pay-rise in nation-wide ‘wage increase campaigns’ that occurred infrequently when the financial condition of the state could afford a general increase (Qian, 1997). Wages generally didn’t increase for very long periods. According to Ding and Warner (2001), the statistics show that the annual growth rate of the real wage-income earned by workers in state enterprises was only 0.3 percent during the period 1953-1977. The situation was even worse during the Cultural Revolution. In the meantime, the wage difference between managers and workers, as well as between intellectual and physical labourers was kept low as egalitarianism was another feature of this system. Since wages were separated from enterprise performance and individual efforts, their role as an effective means of motivating employees and rewarding hard work was seriously undermined (Ding and Warner, 2001).

[5] According to Meng (2000), for example, after 3 years at apprenticeship level, individuals received an increase in wages to level 1 for production workers (8 levels in total), and 3-5 years later, their wage level would automatically increase by a level.

2.3. Labour Market Reforms

2.3.1. Changes of China's Labour Market

According to Kidd and Meng (2001), it has been argued by many Chinese economists and policy makers that the major weaknesses of the pre-reform labour market in China were two-fold: 1) a system of guaranteed life time employment encouraged over-staffing, shirking and low productivity, and 2) the wage grade system rewarded individuals' personal endowments (such as education and work experience) without regard for their actual labour productivity. One issue which is necessary to mention is the consideration of education and work experience by the grade wage system, which appears to be similar to that in a market economy. However, there is an essential ingredient missing as argued by Meng (2000). Human capital theory assumes the existence of a labour market with few imperfections and that each individual decides how much education to obtain to increase his/her marginal productivity and subsequently their future earnings. It is clear that in pre-reform China, the linkage between human capital and labour productivity is broken when enterprises are not able to hire/fire employees; or when employees are not able to choose jobs which may suit their abilities/interest. Generally speaking, there was no mechanism to encourage individuals to work efficiently, thus a high level of human capital stock was not linked to high productivity. The realizations of the low productivity and inefficiency with the pre-reform labour arrangements motivated the labour market reforms inaugurated by the late 1970s. One of the greatest challenges facing China was the development of a well-functioning labour market. Such a labour assignment system should allow employers and employees to choose each other in the light of their needs to obtain actual matches between them.

In 1980, a more flexible labour market strategy began to emerge. Instead of directly being allocated by the government labour bureaus, urban job-seekers were allowed to find work in the state, collective, or newly-recognized private sectors, and enterprises were granted greater flexibility over labour allocation and more autonomy in hiring/dismissal decisions (Brooks and Tao, 2003). In 1986, the introduction of labour contracts allowed state enterprises to hire workers on short-term contracts rather than provide permanent employment. As a result, the share of workers on labour contracts increased, further stimulated by additional reforms in 1994 which the Labour Law of China promulgated, coming into effect on January 1995, which represented a new breakthrough (Ding and Warner, 2001). The share of workers on contracts almost doubled between 1994 and 1997, to about one-third of urban workers (Brooks and Tao, 2003). By the end of 2000, the labour contract system had been implemented in urban enterprises of every description. Restrictions on the mobility of workers between enterprises have also been removed.

Following Deng's southern trip in 1992, with the deepening of the labour market reform, non-state sectors emerged and developed rapidly due to the opening of the economy to private and foreign investors. There was a strong trend to leave the government or state sector, moving to the non-state sectors. As mentioned, the share of total urban employment in state units declined from 62.3% in 1990 to 38.1% in 2000. In the latter part of the 1990s, the government moved forward with aggressive restructuring and privatization of SOEs to resolve the problem of inefficiency, leading to substantial layoffs, retirements, and exits from the labour force (Maurer-Fazio, 1999). Large numbers of redundant employees in SOEs had been laid off, being called *xiagang*[6] workers. An estimated 25 million state-owned enterprise and collective employees were laid off in the period 1998-2002[7]. As part of a reemployment program, the *xiagang* workers could enter reemployment centers to obtain job training and search assistance. As long as they stayed in the reemployment centers, they were provided with minimum income support for up to three years, although they sometimes did not receive that in practice (Yueh, 2004). In March 2005, considering that the structural change of state-owned enterprises had almost been completed and that the peak of *xiagang* workers had already gone, all the reemployment centers of the SOEs were closed within the year. Instead of *xiagang*, the laid-off workers would then be regarded as 'unemployed'.

Beginning with the reform period in the late 1970s and accelerating in the late 1990s, reforms to the household registration (*Hukou*) system have been implemented by the Chinese authorities. Restrictions on obtaining urban residence permits have been gradually relaxed. The extension of the temporary residence system to towns and small cities provided legal channels for rural labourers to move between their rural homes to unskilled jobs in urban areas. Further introduced in 1992, the "blue stamp" *hukou* policy permitted wealthy individuals to buy an "urban" designation by investing large sums of money in urban areas. Apart from this, millions of rural Chinese migrants also migrated

[6] Xiagang workers are one form of laid-off workers who are officially registered as part of their enterprise or work unit, but do not go to work or receive a wage (Yueh, 2004).
[7] Data come from China Statistical Yearbook, 2003.

to urban areas without any formal registration[8]. Despite these reforms, both unregistered migrants and those holding temporary residence permits faced severe limits on their ability to obtain public services such as subsidized health care or education[8]. In 1997, the authorities initiated an experimental program in designated small towns and cities to give migrants an equal chance of being recruited and also allowed those who had either 'a stable job or source of income' or a 'stable place of residence' for over two years to obtain an urban *hukou* (Brooks and Tao, 2003). Since 2001, many provinces and large cities have also begun to allow migrants who satisfy certain criteria to obtain local *hukou* in urban areas. Brooks and Tao (2003) further argued that estimates of the migrant population vary, ranging between 80 million and 150 million between 1990 and 2000[9]. According to official statistics, the urban population share has increased from about 26% in 1990 to 39% in 2002. In addition, of the total growth in the urban population in the 1990s, 55% was due to migration (Chan and Hu, 2003). Migration has been partially the reason behind increasing urban unemployment. Because of Chandrasekhar and Ghosh's (2006) rationale, unemployment in terms of absolute numbers of people has risen steadily since 1985, reaching a high of 7.7 million in 2002.

2.3.2. Wage-setting Reforms of SOEs

Since the beginning of the 1980s, the wage-setting reforms have been implemented to

[8] Detailed discussions see China's Household Registration System: Sustained Reform Needed to Protect China's Rural Migrants, Congressional-Executive Commission on China, October, 2005.

[9] According to Brooks and Tao (2003), the National Bureau of Statistics estimates there were about 80 million permanents migrants (i.e., those living in urban areas for more than six months) between 1990 and 2000. No reliable data are available for the number of temporary migrants, with estimates in the range of 30-75 million.

introduce incentives to work and encourage efficiency (Meng, 2000). These are along two dimensions. Firstly, in 1985, the enterprise's total wage quota system was changed to a 'floating' total wage system. The floating system relates the enterprise's total wage bill to its profitability and economic performance, and it provides profit-oriented incentives (Yueh, 2004). The new system further allows enterprises to retain part of their profits to contribute to their employees' welfare system (Kidd and Meng, 2001). This clearly indicated the increasing level of enterprise autonomy associated with the labour market reforms. In 1993, the minimum wage was set up but there were no unified standards nationwide. The standards of minimum wages shall be fixed and readjusted with comprehensive reference to the following factors: (1) The lowest living costs of localities; (2) Average wage level; (3) Productivity; (4) Situation of employment; and (5) Differences between regions in their levels of economic development[10].

The second dimension involved linking the wages of individuals to their labour productivity within enterprises confined by the total wage budget. In 1994, the new Labour Law was introduced to give enterprises more discretion over wage determination. It is possible for larger income differences to arise, to let employers pay bonuses to more productive workers. The 'wage plus bonus' system was the most frequently adopted and the importance of bonuses has been increased gradually (Meng, 2000). Previously, there had been six components of the wage in enterprises including the basic wage, bonuses, benefits and subsides, overtime wages, supplementary wages and an unnamed component, mainly based on hardship (Yueh, 2004). These categories have been replaced by two

[10] These factors come from the Labour Law of the People's Republic of China, Chapter 5: article 49.

components of wages: fixed wages, including the basic wage, seniority wage, insurance and a housing fund; and bonuses, based on both individual productivity and enterprise profitability (Yueh, 2004). As a result, the share of bonuses in total wages for all enterprises rose from 2.4 percent in 1978 to 23.3 percent in 1993, and then declined to about 16 percent in 1997. In 1994, there seems to be a turning point when the share of fixed wage starts to increase again (Meng, 2000). However, due to high supervision costs, bonuses have often been distributed on an egalitarian basis within work units (Chen, *et al*., 2005). As argued by Ding and Warner (2001), the general wage level of SOEs is still below that of private and foreign-invested enterprises[11]. SOEs are facing difficulties in keeping skilled labour and experienced managers, engineers, and marketing personnel from moving to private and foreign-invested companies in pursuit of more attractive pay packages. This makes it necessary to look at the sectoral differences when estimating returns to education in the fourth chapter.

2.3.3. The Ownership Reform of SOEs

Reforms to SOEs have taken a series of different stages since the late 1970s. According to Guo (2003), the first stage, from 1979 to 1987, centered on the decentralization of management or an expansion of managerial autonomy of SOEs by introducing to enterprises 'mandatory planning reduction', 'profit-retention mechanisms' 'profit-tax

[11] On the other hand, Chen, *et al*. (2005) argued that when non-wage benefits are taken into account, workers in SOEs earn significantly more than workers in urban collectives or domestic private enterprises. Because of the duality of the Chinese economy, foreign-invested enterprises have to pay higher wages to attract skilled workers.

reform[12]', and 'production responsibility systems' under which enterprises could operate more autonomously and efficiently. The second stage, from 1987 to 1992, centered on the separation of ownership and management by introducing a system of 'contracted managerial responsibility', attempting to clarify the responsibilities and benefits between the state and managers. It is clear that such efforts focused on devolving more powers down to enterprise managers and giving them profit incentives without making significant changes to the existing ownership structure (Green, 2003). Since 1992-1993, the third stage reforms centered in the establishment of a 'socialist market economy with Chinese characteristics' by transforming the ownership system into a mixed structure with the public sector dominant and various types of ownership coexistent (Guo, 2003). Starting from the mid-1990s the official reform objective is the creation of a 'modern enterprise system' (MES) with clarified property rights based on separated government and enterprise functions. The central idea of the MES is the commercialization of SOEs by way of transforming them into joint stock companies and limited liability companies (Opper, 2001).

To a large extent, restructuring followed the direction of the official guiding principle: "seizing on the big and letting go the small". Based on this, the state decided to reduce state ownership of small- and medium-size enterprises by mostly selling off or contracting out those running losses. One issue needs to be mentioned, as argued by Chairman Jiang, the aim of 'letting go the small' is not simply to sell them off or abandon them, but to invigorate them. The preferred route for reforming large SOEs is converting

[12] From 1983, profit remittances to the government were replaced by a profit tax. See Lin, Cai and Li (1996).

them into shareholding companies restricted by the central government. Different measures were implemented including mergers, acquisitions, leasing, auction, bankruptcy, formation of shareholding companies or cooperative, and joint ventures. Such reforms have transformed the Chinese ownership system into a mixed structure including state-owned enterprise, collective-owned enterprises, foreign-invested enterprises, private-owned enterprises, individual enterprises, domestic joint-ventures, shareholding companies and others (Chen, *et al*., 2005). It is necessary to emphasize that the state sector has indeed shrunk, however, it continues to be dominant in the national economy. Guo (2003) further argued that China’s ownership reform should be characterized by an increased role of the market and competition, but not by the replacement of public ownership with private ownership as in a capitalist system.

2.4. Regional Inequality

2.4.1. Income Inequality

China’s economic growth has been highly impressive since the 1990s compared to other transitional economies. By 2005, with an average annual economic growth rate of 9.5 percent, real GDP had expanded to 9.7 times that of GDP in 1980. The economic reforms have had a significant impact on income distribution, and at the same time led to a rapid increase in income inequality although it is a historic issue[13]. According to Benjamin, *et al.* (2005), it should come as no surprise that inequality rose as China moved from an

[13] According to Huang, *et al*. (2003), before the Communist government took over mainland China in 1949, the coastal areas had already become more developed than the interior areas.

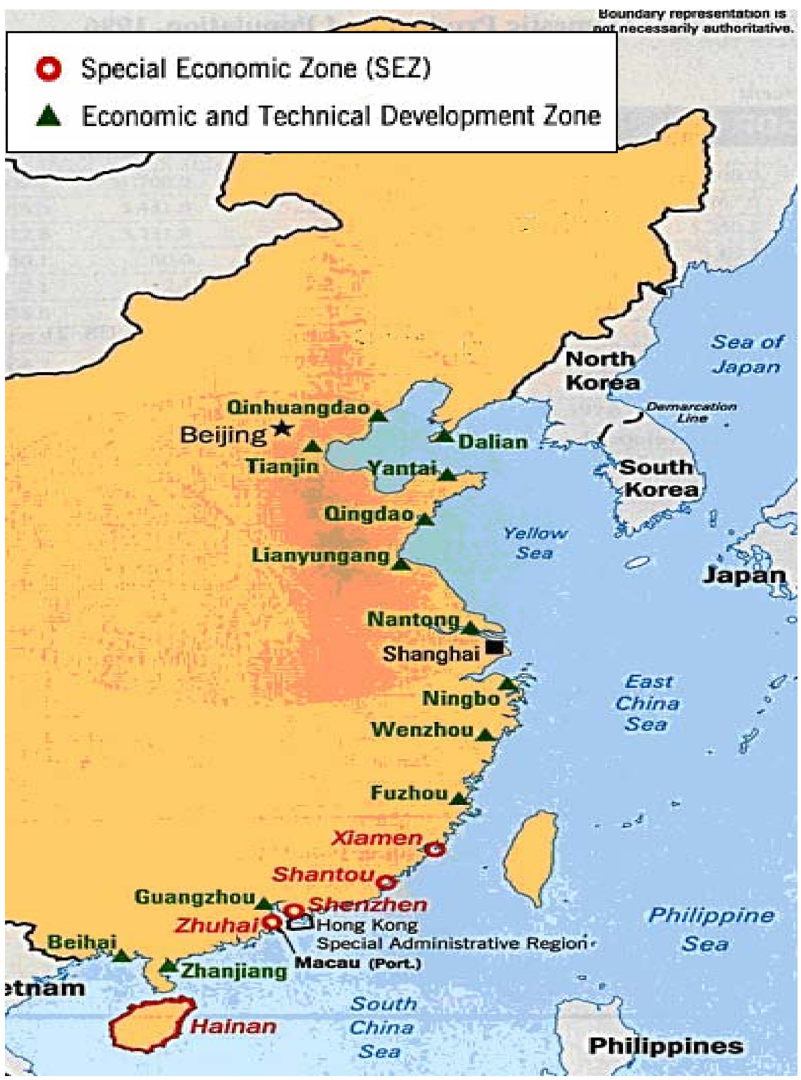
ostensibly egalitarian socialist economy to a more market-oriented one. It is mainly reflected by regional inequality, including intra-rural and intra-urban inequalities, and specifically inter-provincial inequality and rural-urban inequality which we are going to generally describe in this section. A key dimension of inter-provincial inequality is that between inland and coastal provinces. We will begin with this. The China's economic reforms were introduced gradually towards making the economy more market-oriented and opening up to the outside world. Since Deng, the chief architect of China's economic reforms, believed that China could not achieve rapid economic growth in all regions simultaneously (Jones, *et al.*, 2003). The new strategies were designed to allow a portion of people and areas to grow rich first. In other words, the economic reforms shifted the focus from encouraging full regional development to developing only the coastal regions in the southern and eastern parts based on their readily accessible geographic location (Huang, *et al.*, 2003).

The central government further established the special economic zones (SEZs) of Shenzhen, Zhuhai, Shantou, and Xiamen (1980) (see Figure 2-1), 14 open coastal cities (1984), Hainan island (1988), Shanghai Pudong District (1990), free trade zones in coastal cities (1993), open border cities and open free trade zones such as Yangtze Delta and the Pearl River (Jones, *et al.*, 2003). The introduction of foreign direct investment (FDI) and technology concentrated on these limited regions together with granted preferential policies[14] by central government, which has brought rapid economic growth to them (Okushima and Uchimura, 2005). During the period of 1985 to 2000, more than 85% of the cumulative FDI was located in the coastal (or eastern) regions (Wei, 2003/2).

[14] For detailed discussions on these 'preferential policies' see Jones, *et al*., 2003.

The emphasis on coastal development has led to a rapid rise in the income inequality in China. For instance, in 2000, per capita GDP and the total GDP of Jiangsu province, one of the eastern coastal provinces, were 4.4 times and 8.6 times[15] larger respectively than Guizhou province, one of the western provinces where economic development has been relative slow. See Table 2-1.

Figure 2-1 China Special Economic Zones



Source: University of Texas Libraries, map collection.

---

15 Data come from China Statistical Yearbook, 2001.

Table 2-1 Gross Domestic Product (GDP) and Per Capita GDP by Regions

Eastern	GDP (100M yuan)			Central	GDP			Western	GDP		
	1996	2000	2005		1996	2000	2005		1996	2000	2005
Beijing	1615.73	2478.76	6886.31	**Jilin**	1337.16	1821.19	3620.27	**Shaanxi**	1175.38	1643.81	3675.66
Shanghai	2902.20	4551.15	9154.18	**Heilongjiang**	2402.58	3253.00	5511.50	**Sichuan**	4215.00	4010.25	7385.11
Tianjin	1102.40	1639.36	3697.62	**Inner Mongolia**	984.78	1401.01	3895.55	**Guizhou**	719.83	993.53	1979.06
Hebei	3452.97	5088.96	10096.11	**Anhui**	2339.25	3038.24	5375.12	**Yunnan**	1491.62	1955.09	3472.89
Liaoning	3157.69	4669.06	8009.01	**Jiangxi**	1517.26	2003.07	4056.76	**Gansu**	714.18	983.36	1933.98
Jiangsu	6004.21	8582.73	18305.66	**Shanxi**	1305.50	1660.92	4179.52	**Qinghai**	183.57	263.59	543.32
Zhejiang	4146.06	6036.34	13437.85	**Henan**	3683.41	5137.66	10587.42	**Ningxia**	193.62	265.57	606.10
Fujian	2606.92	3920.07	6568.93	**Hubei**	2970.20	4276.32	6520.14	**Xinjiang**	912.15	1364.36	2604.19
Shandong	5960.42	8542.44	18516.87	**Hunan**	2647.16	3691.88	6511.34	**Tibet**	64.76	117.46	251.21
Guangdong	6519.14	9662.23	22366.54								
Guangxi	1869.62	2050.14	4075.75								
Hainan	389.53	518.48	894.57								
	Per capita GDP (yuan)				**Per capita GDP**				**Per capita GDP**		
Beijing	15044	22460	45444	**Jilin**	5163	6847	13348	**Shaanxi**	3313	5137	9899
Shanghai	22275	34547	51474	**Heilongjiang**	6468	8562	14434	**Sichuan**	3763	4784	9060
Tianjin	12270	17993	35783	**Inner Mongolia**	4259	5872	16331	**Guizhou**	2093	2662	5052
Hebei	5345	7663	14782	**Anhui**	3881	4867	8675	**Yunnan**	3715	4637	7835
Liaoning	7730	11226	18983	**Jiangxi**	3715	4581	9440	**Gansu**	2901	3838	7477
Jiangsu	8447	11773	24560	**Shanxi**	4220	4549	12495	**Qinghai**	3748	5087	10045
Zhejiang	9455	13461	27703	**Henan**	4032	5444	11346	**Ningxia**	3731	4839	10239
Fujian	8136	11601	18646	**Hubei**	5122	7188	11431	**Xinjiang**	5167	7470	13108
Shandong	6834	9555	20096	**Hunan**	4130	5639	10426	**Tibet**	2732	4559	9114
Guangdong	9513	12885	24435								
Guangxi	4081	4319	8788								
Hainan	5500	6894	10871								

Source: China Statistical Yearbook, various years.

The two largest cities of Beijing and Shanghai account for a large part of the variation in regional incomes. These two large metropoli, which enjoy a high level of industrialization and a large percent of their population living within 100km off the coast or on navigable waters, which enabled them to reap the full benefits of public infrastructure expansions, export promotion and substantial FDI inflows (Chandrasekhar and Ghosh, 2006). For example, the difference in per capita GDP between Shanghai and Guizhou increased from 7.3 times in 1990 to 12 times in 2000[16]. Given that the coastal economy has largely expanded, the central government started to show concern for the development of interior regions in the 1990s. The plan of 'the Great Development of the Western Area', which has been promoted since 1999, may represent the government's *first real determination* to develop its vast interior (Huang, *et al.*, 2003). The recent strategies have involved funding new infrastructure projects and increasing fiscal transfers to these areas (Chandrasekhar and Ghosh, 2006). It is hard to actually assess this strategic shift because the development of the western regions is 'a systematic project and a long-term task, which may take the efforts of several generations'[17]. In addition, it must be recognized that almost all large economies exhibit substantial regional differences. This is not specially a Chinese phenomenon.

Han (1998) has argued that the urban-rural inequality has been the biggest contributor to the whole problem of inequality in China. As mentioned, the increases in inequality are related to the economic growth process. In rural areas, the non-farming income opportunities, which significantly contribute to the inequality, are concentrated in a few

[16] Data come from Shijie Ribao, January, 2000.
[17] Borrowing the words of Chinese leaders in *Beijing Review*, April, 2000.

areas, whilst rich urban areas, including for example urban municipalities, have grown more rapidly than others. This gap in urban areas is exacerbated by Chinese fiscal policies, which allow these richer areas to retain their higher revenues for local use and benefits. According to Sicular, *et al.* (2005), China's urban-rural gap is not uniform regionally, it is higher in the western regions and comparatively lower in the central and eastern regions. For example, as argued by Benjamin *et al.* (2005), in the more dynamic coastal provinces, more rapid job growth in the non-state sector helped reduce urban-rural differentials by fostering more rapid rural income growth. The level of urban-rural inequality is significant in China, and the income gap between urban and rural areas rose steadily from 1979 to 1994. Chandrasekhar and Ghosh (2006) argued that this inequality has largely accounted for the growing gap between wages in agriculture which is dominant in rural areas and wages in industry and services which predominate in urban areas. There was a short period of declining differences between 1995 and 1997, based on the increased price of agricultural products, introduced by the government in 1995 (Lu and Chen, 2006). Since 1997, inequality has increased once again as the price of agricultural products has fallen. Apart from these factors, restrictions on labour movement have been another important factor behind urban-rural inequality in China.

The urban-rural income gap is often attributed to policies that inhibit labour mobility including, for example, discouraging relocation and employment of rural migrants in urban areas which is specifically restricted by the household registration system as mentioned before. Lu and Chen (2006) argued that the outflow of rural labourers increased the income of rural residents because migrant workers send money back to their hometowns. From the 1980s onwards, some rural migrants have been granted temporary residence permits in urban regions. However, the larger part of the migrant population did not qualify for those permits, and had to remain in the informal sectors without access to the public utilities (Chandrasekhar and Ghosh, 2006). Although recent social security system reform has shifted the responsibility from the SOEs and aimed to set up a unified system, it is clear there still exists a great variability in coverage of workers and the rural migrants are mostly left out of system. They are discriminated against in terms of education, housing, public service, pensions and health insurance[18] compared to urban residents. Furthermore, to protect urban residents, many large and medium-size cities limited rural labour's entry into good jobs by administrative means[19] (Lu and Chen, 2006). These significant discriminations in social welfare significantly increased the urban-rural income differentials. Although the household registration system relaxed recently, huge challenges remain. For example, it is still noteworthy that richer rural inhabitants tend to have a higher possibility of obtaining urban household registrations than poorer inhabitants[20].

2.4.2. Inequality of Health Care

China has experienced unprecedented economic growth, which has been accompanied by dramatic increases in income inequality as mentioned above. Relatively little discussion

[18] On health insurance, Han (1998) argued that although some rural residents who work for country or township governments have employer-covered health insurance, the quality and availability in urban areas far exceeded that in rural areas.
[19] For example, according to Chandrasekhar and Ghosh (2006), rural migrant workers in cities often face discrimination as a result of local regulations, being charged various 'administration fees'.
[20] For detailed discussions see Lu and Chen, 2006.

is available on inequality in other dimensions of human development related to this income inequality, such as health care and education, which we are going to generally describe in this and the following section. Before the economic reforms, almost all urban residents of working age had guaranteed jobs in the SOEs and COEs. These jobs provided urban residents with benefits such as free or subsidized education, housing and health care. Two primary components of health insurance schemes operated in urban China in the pre-reform era. The Government Insurance Scheme[21] (GIS), which covered government employees, retirees, disabled veterans, and university teachers, staff and students; and the Labour Insurance Scheme (LIS), which covered state enterprise employees, retirees and their dependents (Dong, 2001/2). The former was financed directly by government budgets, and the latter was financially supported by the welfare funds of enterprises. These two schemes provided comprehensive health care benefits and services based on a person's place of employment.

Health care and education attainment for rural residents were much worse compared to urban ones, due to the government's urban-biased policy (Zhang and Kanbur, 2005). The Cooperative Medical System (CMS) was established primarily in rural areas, together with a three-tier (county, commune and production brigade) health delivery system, which guaranteed the majority of rural population access to essential health care regardless of their economic situation. By 1979 it is estimated that 80-90% of the rural population were covered by some form of CMEs (Zhu and Zhang, 1987). China's health care system has undertaken great changes since the start of economic reforms for both urban and rural areas specifically after the 1980s, which has led to increasing inequality

[21] In China, this plan is more often called 'Free Medical Care'.

of health care coverage, cost and financing.

China's reforms have led to significant changes in health care coverage and costs with the collapse of the CMS in the rural areas and the erosion of the LIS in the urban areas (Wu, 2003). Because the GIS is financed by each provincial government, the coverage of GIS in urban areas differs by province as well, with eastern provinces more capable of providing health insurance coverage to government employees than provincial governments in the western and central regions (Grogan, 1995). However, the major difference with respect to health insurance coverage is between urban and rural areas. In the urban areas, some urban programs have introduced cost-containment incentives to insurance programs, including co-payments and fixed monthly payments, which recipients can use for medical care or keep if they are not ill (Shi, 1993; Guo, 2003). However, according to Zhao (2006), in large and medium-sized cities, respondents having no health care coverage accounted for some 40 percent of the population in 2003, and in rural areas, proportions covered by the resurrected CMS or other health insurance schemes were generally small, and more than 80 percent had no health care coverage at all. An increasing number of people have to bear the full costs of medical care[22], the majority of rural residents, whose income is on average less than half of the urban residents[23] have to pay for health care out-of pocket, and those who can not afford to do so are denied access to care. Feng, *et al.* (1995) argued that ill health was one of the major contributors to the poverty of families in the poorer areas, and that almost half of those who needed medical care could not afford it. It is necessary to further mention that

[22] According to Zhang and Kanbur (2005), the share of out-of-pocket expense in medical care for China as a whole increased from 16% in 1980 to 61% in 2001.
[23] For detailed discussions see Shi, 1993.

rural-to-urban migrants do typically earn more than the average rural residents, however it is unclear whether they actually have more health care purchasing power because as temporary migrant workers, they are not privileged to any urban health insurance programs (Grogan, 1995).

Decentralization has been a central feature of fiscal reforms in the early 1980s. In this environment each province and lower levels of government are responsible for their own health care service, because of regional differences in prosperity, thus accentuating the already uneven distribution of health care services between rural and urban areas (Shi, 1993). The proportion of the health care expenditure directly borne by government has declined sharply, and it has been specifically inadequate in less developed rural areas. In 2000, the per capita rural health spending was merely 12 yuan, only 27.6% of the urban spending. Of the health expenses for China, 70% of the rural population accounted for only 33% of the country's total[24]. It is still early to close the discussion on whether urban-rural inequality in health care service leads to differential health outcomes, but it is clear that the urban population enjoys better health status than those who live in rural areas. For example, in 2000, urban residents have lower mortality rates than their counterparts in rural areas, and, further, people living in large and medium-sized cities live much longer than those in less developed rural areas (Zhao, 2006). It is clear that further reforms on health care system are still necessary. More effort needs to be made by the Chinese government based on a universal and compulsory health financing system, with regard to increasing health care coverage among the population and controlling spiraling increases in medical costs (Shi, 1993; Zhao, 2006).

[24] Data come from the Red Flag Manuscripts Magazine, China, 2003.

2.4.3. Inequality of Education Attainment

People in China have traditionally placed a high value on education, and the reason may be found not only in Confucian attitudes, but also in the potential benefits (Knight and Shi, 1996). Education is an important determinant of not only an individual's income, but also his/her general well-being including factors such as improvements in health and nutrition. The system of education in China is divided into three categories: basic education, higher education, and adult education. Basic education includes primary education (normally six years) and regular secondary education. Secondary education is divided into academic secondary education (normally three years of lower and three years of upper middle school) and specialized/vocational/technical secondary education. Higher education at the undergraduate level includes two-and three-year junior colleges, four-year colleges, and some universities offering programs in both academic and vocational subjects. Many colleges and universities also offer graduate programs leading to the master's or Ph.D. degree[25]. Since the issuing of the Law of Compulsory Education in 1986, governments at all levels have aimed to promote nine-year compulsory education (six years of primary education and three years of lower middle school) throughout China. After almost twenty years, this goal has not been totally accomplished in rural and some remote areas, however, it has made remarkable strides in the elimination of illiteracy. For example, the illiteracy rate in 1990 was approximately 21% of the total population, whereas in 2002 this figure was reduced to around 10%[26]. At the same time, higher education in China has experienced a massive expansion, especially over the last two

[25] Education system in China, see China Education and Research Network, www.edu.cn.
[26] Data come from China Population Census, 1990; 2002.

decades, and the majority of high school graduates in China go on to attend college. According to the China Statistic Yearbook 2003, the proportion of upper middle school graduates entering into an institution of higher learning was 49.9% in 1996, and it has risen to 83.5%[27] in 2002.

According to Qian and Smyth (2005), on a larger scale, education plays an important role in contributing to social equity. Similar to health care, disparities exist in educational attainment, and specifically, gender inequality in education persists in China. Rong and Shi (2001) argued that the high level of female illiteracy and the underprivileged situation of girls with regards to their access to and retention in schools is one of the greatest challenges facing China. Table 2-2 shows the percentage of education attainment by educational level and gender in 1990 and 2000. It is clear that the percentage of illiteracy for women has substantially declined from 1990 to 2000, however, it was still 9% higher than for men in 2000. In addition, women accounted for 72.7% of the illiterate (and semi-illiterate) population[28], and this related to two points. Firstly, the high proportion of illiteracy for women is largely focused on the older population aged over 45. For people aged less than 45 (15-44), the percentage of illiteracy for women and men was 3.2% and 1.1% respectively in 2000[29]. Secondly, as shown in Figure 2-2, women (and men) in rural areas accounted for a large proportion of the illiterate population. This is consistent with the fact that girls (and even boys) in poor rural areas have less opportunity to attend

[27] Data on the proportion of upper middle school graduates entering into institutions of higher learning refer to the ratio between new entrants into regular institutions of higher learning and graduates of upper middle schools (China Statistic Yearbook, 2003).
[28] These figures relate to the whole population, many of whom left school 20 or more years ago. Even if today the system is improving, it will still take decades to remove inequalities from the population as a whole.
[29] China Population Census, 2000.

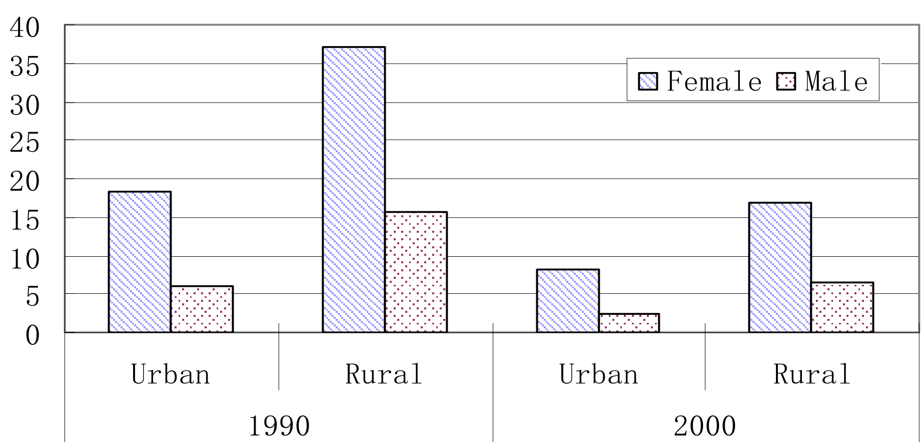
school due to the impacts of a less developed economy and traditional conventions.

Table 2-2 The Percentage of Education Attainment by Level and Gender (Population Aged 15 and over)

Level	1990 Female	1990 Male	2000 Female	2000 Male
Percentage Distribution (%)				
Illiteracy	31.9	13.0	13.9	4.9
Primary School	33.8	35.3	33.6	28.4
Lower Middle School	24.3	36.0	36.0	44.6
Upper and Specialized Middle School	8.8	13.1	12.8	16.4
College and over	1.2	2.6	3.7	5.6
Total	100.0	100.0	100.0	100.0
Gender Distribution (%)				
Illiteracy	70.1	29.9	72.7	27.3
Primary School	47.7	52.3	52.7	47.3
Lower Middle School	39.0	61.0	43.2	56.8
Upper and specialized Middle School	39.1	60.9	42.4	57.6
College and over	30.3	69.7	38.3	61.7
Total	48.8	51.2	49.0	51.0
Numbers in Millions	398.6	419.0	469.1	489.0

Source: China Population Census in 1990 and 2000.

Figure 2-2 Illiteracy Rate by Urban/Rural Areas in 1990 and 2000 (Population Aged 15 and over)



Source: China Population Census in 1990 and 2000.

Apart from the inequality between genders, inequality between the regions is mainly based on the quality of education provision and government expenditures, the contrasts of which stand out prominently in contemporary China. Data from the late 1990s and the year 2000 show that economically advantaged regions and urban areas enjoy substantial advantages in educational provision (Zhang and Kanbur, 2005). For example, from 1990 to 1998 the pre-primary level in the rural areas witnessed no increase in the total number of teachers and school facilities in spite of the increased enrolment rate, however the urban areas saw an increase in both the number of school facilities and the number of teachers as well as in enrolments[30]. At the primary level, sixteen (mainly coastal) provinces had a 99.5% enrolment rate while a number of inland provinces (more heavily populated by minority groups) had markedly lower rates[30]. There is considerable administrative and fiscal decentralization in China. The structure of educational financing has undergone a fundamental change from a centralized system with a narrow revenue base to a decentralized system with a much more diversified revenue base since the early 1980s (Qian and Smyth, 2005). Educational expenditures rose while the government's share in educational expenditures dropped in the 1980s and 1990s (Tsang, 2000). Educational expenditures as a percentage of GDP have remained at a fairly low level (less than 2.4%) throughout the 1990s (Rong and Shi, 2001). According to Knight and Shi (1996), when education is publicly funded and funding is regionally decentralized, it is reasonable to expect that a prosperous region can afford to devote more resources to education, whereas a poor region can afford less. Thus the increasing emphasis on local self-reliance without cross-subsidization from richer to poorer regions, made regional inequality with respect to both the quantity and quality of educational provision and

[30] See *Basic Education*, in United Nations in China, www.unchina.org.

attainment more significant.

2.5. Conclusion

This chapter generally reviewed the economic reforms and the existing regional inequalities during China's transition from a centrally-planned to a market economy. Even through the overall performance of China's transition is very remarkable, the process is not complete yet. For example, as argued by Guo (2003), many SOEs after being transformed into shareholding forms continue to act and function in the same way as they did before, i.e., the old problems are still there, such as redundant employees, bad debts and all kinds of burdens and charges on enterprises. SOEs are owned by the state, but the control rights are divided between government and enterprise managers. Managers are not owners or pure shareholders, thus they can not make their decisions completely based on the enterprise's economic performance, and recruit or dismiss employees on their abilities and productivity. As Opper (2001) concluded, the more diluted property rights are, the more enterprises are subject to political interventionism by the local and state administration, and the worse the structural performance will be.

According to Lardy (2002), the absolute living standard has gone up enormously, although China has undergone a rapid increase in income inequality, and fairly rapid economic growth is important in making inequality more manageable. China is a member of the World Trade Origination (WTO), joining in December 2001. China's entry into the WTO makes its door even more open to the world. It is necessary to continue with further reforms and progress based on introducing greater competition, and these reforms and the accompanying globalization of the Chinese economy have to proceed at an appropriate speed to be consistent with a foreseen dramatic reduction in employment, specifically in the state sector. Some areas need to be given specific attention for further development, for example, the improvement of technology in agriculture, environmental protection and the social welfare systems, including unemployment insurance schemes and other social safety nets. Furthermore, as the Chinese economy becomes a more mature market economy and more integrated with the world economy, according to Lin (2004), it is crucial for the continuous growth of that economy to establish a transparent legal system that protects property rights so as to encourage innovations, technological changes, and domestic as well as foreign investments in these economies.

Chapter 3 Theoretical Framework and Empirical Background

The importance of education in economic development has been emphasized in economic theory for many years. The concept of human capital is central to much of the research in the economics of education. Prior to this research, economic theories of capital and investment tended to concentrate on investment in physical capital, thus the development of human capital theory broadened the concept of capital and recognized that human skills were like any other physical capital in that they could be augmented through investment. Education is one of the most important aspects of investment in human capital, and specifically, we focus on education as a private investment decision. Following the introduction, this chapter discusses the theoretical and empirical issues relevant to educational returns corresponding to earnings, health and subjective well-being. We will begin with the human capital, focusing on education, and further expand to health capital and social capital.

To estimate the impact of education in the following empirical chapters, it is necessary to start with how education is measured. Historically, education generally has been represented in the empirical studies by the number of years or grades that an individual has completed in the formal system. The term which is commonly used is 'educational attainment' (Behrman, *et al.*, 1997). Education is correlated with many outcomes, and most of these correlations used educational attainment to represent education. The reason for basing this on 'formal education' is, as argued by Behrman (1997), that it is a major form of education in most societies, and further it indicates that more information is gained than in most other forms of education. For example, training, as one form of education, is directly focused on the productivity effects of education rather than non-labour market benefits.

There exists an argument in the literature based on the comparison of years of education completed and the highest level of education obtained. The idea is reflected in the so-called 'sheepskin' effect where qualifications have a discounted return that exceeds the return that could otherwise be gained from the number of years spent acquiring them so that there are discontinuities in the returns to education at points associated with acquiring qualifications (Chevalier, *et al.*, 2004). For example, there might be a wage premium over the average return for fulfilling a particular year of education, such as the final year of college. Observing the non-linear wage returns to schooling years has been the basis of traditional tests for sheepskin effects (Gibson, 2000). However, the stronger test would be to compare the earnings of those who have a degree and those who do not, conditional on both groups having the same number of years of schooling (Park, 1999). It is clear that such a comparison is possible based on the available information of both years of education and qualifications received. In the thesis, we use the number of years completed in formal system as a main indicator of education as this is consistent with much other research.

As suggested by Behrman, *et al.* (1997), education might generate benefits in three ways: by changing individuals' preference, by changing the constraints individuals face, or by augmenting the knowledge or information on which individuals base their behavior. As a result, the earnings of more educated people are almost always well above average; more educated people are more likely to have a healthy life style and thus less likely to be obese. However, these correlations do not necessarily imply causality. Correlation may be due in part to causal effects of education, but it also may reflect reverse causality or the fact that both education and outcomes respond to other determinants, such as innate ability, inherent health and family background, which make the causal explanation difficult. Therefore, together with introducing the theoretical foundations, this chapter pays particular attention to a number of the most important difficulties.

This chapter is organized as follows: in the next section, we present the human capital framework: we start with the typical human capital theory developed by Becker (1964), followed by Mincer's (1974) specification: the benchmark model for empirical estimation of the returns to education. Specifically, we also focus on the possible difficulties surrounding this form of analysis. Section 2 discusses health capital based on Grossman's (1972) model, with the emphasis on the possible linkages between education and health.

In section 3, we introduce the dimensions and measurement of social capital and generally discuss the linkage between some social capital variables and subjective well-being. Finally we conclude this chapter.

3.1. The Human Capital Framework

3.1.1. Internal Rate of Return on Basic Theoretic Model of Educational Investment

Human capital theory provides a framework for analyzing educational investment. The acquisition of education generates costs, and education is assumed to raise an individual's productivity and consequently to pay off in terms of future earnings. According to the traditional method of calculating rates of return to investment in education, which is known as cost-benefit analysis, the analysis must commence with the calculation of all costs and benefits involved in this investment. It is suggested that although in theory it is straightforward, the actual measurement of costs, and especially benefits is difficult to obtain. Based on this, economists have tended to concentrate on the relatively hard evidence relating to *benefits* that exists in most countries, which is known as the age-earnings profile, that those people with higher levels of education on average receive higher incomes throughout their working lives than people with lower levels of education (Hough, 1994). In Becker's (1994) words, education (schooling) would steepen the age-earning profile. It has therefore seemed reasonable to look at the earnings-stream differentials to measure the benefits from received education.

Two related concepts which are often used in this context are the net present value (NPV) and the internal rate of return (IRR). The NPV of an investment is the difference between discounted benefits and discounted costs based on a preselected discount rate, and the IRR is the discount rate which equalizes the discounted benefits and discounted costs. These two methods are based on the same principles and in many cases they give equivalent answers (Cohn and Geske, 1990).

The following discussion draws heavily on Becker (1964)'s methodology by using costs of education and economic returns on education investment as inputs to derive an internal rate of return. Let Y be an activity providing a person, entering at a particular age, with a net earning stream of Y_0, Y_1, …, Y_n, and X another activity providing a net earning stream of X_0, X_1, …, X_n. Based on the present value of V(Y) and V(X), the net present value of the gain from choosing Y would be given by:

$$V(Y) - V(X) = \sum_{j=0}^{n} \frac{Y_j - X_j}{(1+i)^j} \tag{3.1}$$

Where i is the market discount rate, and assumed to be same in each period. The term ‘activity’ is used following Becker to indicate that any kind of investment in human capital is permitted, although we specifically focus on education. ‘Net’ earnings mean ‘gross’ earnings during any period minus tuition costs during the same period. Those include the sum of both monetary and the monetary equivalent of non-monetary earnings[31]. Equation (3.1) can be reformulated to bring out the relation between costs and benefits. The cost of investing in human capital equals the net earnings foregone by choosing to invest rather than choosing an activity requiring no investment. Suppose *Y* requires an investment only in the initial period[32] and *X* does not require any. The cost of choosing *Y* rather than *X* is simply the difference between their net earnings in the initial period, and the total benefit would be the present value of the differences between the net earnings in later periods (Becker, 1994). If $C = X_0 - Y_0$, $k_j = Y_j - X_j$, and R measures the total benefit, Thus:

$$\sum_{j=1}^{n} \frac{Y_j - X_j}{(1+i)^j} + (Y_0 - X_0) = \sum_{j=1}^{n} \frac{k_j}{(1+i)^j} - C = R - C \tag{3.2}$$

Individuals are assumed to choose their educational attainment in order to maximize the net present value of the stream of future incomes, and the optimal amount of education is

---

[31] Becker (1994) argued that the term “investment in human capital” is not restricted to monetary costs and returns. It applies independently of the division of earnings into monetary and psychic components. For example, it applies to health which has a large psychic component. Later in this thesis we develop this analysis further.

[32] The investment in human capital is distributed over many periods, and this is to simply the analysis. For further discussions see Becker (1994).

obtained when the present value of returns equals the present value of costs in the same period (*R* = *C)*. This relation between costs and benefits can be expressed by the internal rate of return, *r*, which is defined implicitly by the equation:

$$C = \sum_{j=1}^{n} \frac{k_j}{(1+r)^j} \qquad (3.3)$$

When the returns are the same in each period, k_j = k, the equation (3.3) is considerably simplified and becomes:

$$C = \frac{k}{r}[1-(1+r)^{-n}] \qquad (3.4)$$

As people live longer, *n* tends to infinity, where $(1+r)^{-n}$ tends towards zero. Thus the equation (3.4) becomes:

$$C = \frac{k}{r} \qquad (3.5)$$

The aforementioned discussion indicates that if investment in education is restricted to a single known period, the cost and rate of return are easily determined from information on net earnings. It is suggested in the literature that the sense of *n* in equation (3.3) can be implied by T and t: the time horizon T, which is fixed by the retirement age or death; the choice variable t is the individuals to choose in how many years they invest in education. The returns are better if the educational investment is made earlier, which corresponds to the age-earnings profile. Since individuals invest early in their lives not only by means of education, but also by acquiring information about the labour market in their future career, thus later in their life they reap the returns in the form of higher earnings. Furthermore, foregone earnings are a major component of the costs of education, thus older people invest less in education since they are likely to have higher earnings. More detailed discussions by Harmon, *et al.* (2003) where k_j = k = $W_s - W_{s-1}$, and C is expressed by adding the earnings of a graduate of school level W_{s-1} to the direct cost of school level C_s, where $C = W_{s-1} + C_s$, thus:

$$\frac{W_s - W_{s-1}}{r_s} = W_{s-1} + C_s \tag{3.6}$$

The optimal investment decision would imply that individuals invest in the s^{th} year of schooling when $r_s > i$. If C_s is sufficiently small, thus:

$$r_s \approx \frac{W_s - W_{s-1}}{W_{s-1}} \approx \log W_s - \log W_{s-1} \tag{3.7}$$

which shows that the return to the s^{th} of schooling is approximately the difference in log wages between leaving at times *s* and at (s-1)[33].

The empirical approximation of the aforementioned theoretical framework is the familiar functional form of the earnings equation developed by Mincer (1974). Mincer has gone further in analyzing the effect of age on earnings in terms of the acquisition of experience. The Mincerian equation has been proved to be a powerful tool to explain what determines earnings and to compare the relative importance of education and other factors, such as gender, race, region, etc. in determining earnings. Since it is relatively quick and easy to compute, and does not include any specific reference to direct educational costs, although it does incorporate earnings forgone, it has become the benchmark model for the empirical estimation of the returns to education.

3.1.2. The Mincer Specification

The concept of an earnings function was used by Mincer (1974) to explain the pattern of individual earnings. The benchmark model for the development of empirical estimation of the returns to education is the key relationship which involves regressing the natural logarithm of earnings (Y) against educational attainment (edu) and working experience (exp). This is expressed as follows:

---

[33] For detailed discussions see Harmon, *et al*. (2003): "The Returns to Education: Microeconomics".

$$\mathrm{Ln}(Y_i) = \beta_0 + \beta_1 edu + \beta_2 exp + \beta_3 exp^2 + \varepsilon_i \qquad (3.8)$$

where ε_i is a stochastic error term, often it is assumed that $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$. This form of earnings equation is the simple one which assumes that earnings are determined only by two factors: years of schooling and work experience. In the absence of direct information on work experience, Mincer mentioned the use of 'potential experience': the number of years an individual of age A could have worked, assuming he started to school at age 6, finished E years of schooling in exactly E years, and began to work thereafter: potential experience = A – E – 6 (Card, 1999). Furthermore, it is also possible to estimate more complex earning functions that incorporate additional variables including age, gender, occupation, urban or rural location, and some mostly unobserved variables, such as personality, ability and motivation, etc. Thus the equation (3.8) can be often expressed by:

$$\mathrm{Ln}(Y_i) = X_i\beta + \gamma edu_i + \varepsilon_i \qquad (3.9)$$

Where Y_i is an earning measure for an individual i, and X_i is a set of conditioning variables mentioned above. edu_i is the number of years schooling, and the parameter γ indicates the estimated increase in earnings from one additional year of schooling. ε_i is the error term which captures all relevant factors in earnings determination that are not captured by observable variables. In order for this interpretation to be valid, we need to impose certain conditions on the error term. We assume $E(X_i \varepsilon_i) = 0$ and $E(edu_i \varepsilon_i) = 0$, and make the standard hypothesis that $E(\varepsilon_i) = 0$. These assumptions allow us to interpret this equation as the conditional expectation of $Ln(Y_i)$ given X_i and edu_i. A large literature has grown up around the Mincerian model using a variety of variables. Two of the most comprehensive surveys are presented by Psacharopoulos (1985, 1994) covering the results of estimating the returns to human capital for over sixty countries. According to these results, educational returns for an additional year of schooling are positive and range from 5% in developed countries to as high as 29% in developing countries.

3.1.3. Endogeneity of Education based on Ability Bias[34] in the Mincer Model

If it is not the case as mentioned above that $E(edu_i \varepsilon_i) = 0$, although we can still interpret the equation as the conditional expectation of $Ln(Y_i)$ given X_i and edu_i, it is not valid to interpret γ as the causal effect of an unbiased estimate of education on earnings since education is endogenous with respect to the causal effect γ. Broadly speaking, a regressor is endogenous when it is correlated with the error term. If any one regressor is

[34] According to Cameron and Trivedi (2005), leading examples of endogeneity, dealt with extensively in empirical model settings, include simultaneous equations bias, omitted variable bias, sample selection bias and measurement error bias. Endogeneity is quite likely to occur when cross-section observational data are used, and economists are very concerned with this complication. It had been suggested that there exist other possibilities apart from ability bias based on the endogeneity of education. For the purpose of this thesis the focus will be on the ability bias.

endogenous then in general the estimates of all regression parameters are inconsistent and this applies to both linear and non-linear models (Cameron and Trivedi, 2005). Human capital theory treats education as investment by individuals in themselves, and the coefficient of education (γ) in equation (3.9) is broadly interpreted as a measure of returns to human capital. It is clear that education is not randomly assigned; rather, it depends on the choices made by individuals, families, even communities. The regression (3.9) is then a regression of one endogenous variable, Y, on another, edu, and it does not measure in an unbiased manner the causal impact of an exogenous change in education. Thus the observed difference in earnings between individuals with different years of education may over-state or under-state the causal effect of education on earnings[35].

Education is indeed an endogenous choice variable in the underlying human capital theory. Specifically, in the Mincer specification the error term captures unobservable individual effects and these individual factors may also influence the education decision, and therefore induce a correlation between education and the error term in the earnings function. A common example is unobserved ability. There have been a number of approaches to deal with this problem. Firstly, measures of ability have been incorporated to proxy for unobserved effects. The use is documented in the literature to obtain a good measure/proxy approximating the impact of ability, such as intelligence quotient (IQ)

[35] For example, if increasing ability (increasing μ_i in equation (3.10)) increases schooling but also independently increases earnings, then $E(\mu_i\varepsilon_i) > 0$ and the standard OLS estimator will over-estimate γ as the causal effect of education as it will also pick up the effect of characteristics in the error term (such as ability) which makes individuals earn more.

scores, and include it as a direct measure of ability in the earning equation. This inclusion should reduce the estimated education coefficient if it acts as a proxy for ability, so that the coefficient on education then captures the effect of education alone. The usual finding is that the coefficient on education is not strongly affected by the inclusion of additional right-hand-side variables (Weiss, 1995).

Secondly, there exists an approach dealing directly with the education-earnings relationship in a two-equation system by exploiting instrumental variables. To understand more clearly how this approach has been derived, it is helpful to posit the following equation together with equation (3.9):

$$Edu_i = X_i\lambda + \mu_i \tag{3.10}$$

where X_i includes conditioning variables that directly influence both education and earnings, and μ_i is the error term. We assume $E(edu_i\,\varepsilon_i) = 0$ which is to say $E((X_i\lambda + \mu_i)\,\varepsilon_i) = 0$, therefore, $E(X_i\lambda\varepsilon_i + \mu_i\varepsilon_i) = 0$, and therefore $\lambda E(X_i\varepsilon_i) + E(\mu_i\varepsilon_i) = 0$. Assuming that the error term ε_i is orthogonal to the X variables, this formal part $\lambda E(X_i\varepsilon_i)$ will be zero, for the expression to equal zero requires that $E(\mu_i\varepsilon_i) = 0$. If we expect that factors affecting education will also affect earnings, this condition will not hold. To put this differently, estimation of equation (3.9) yields inconsistent estimate of γ when the pair of residuals ε_i

and μ_i are correlated (e.g., in this case the unmeasured ability is a confounding variable[36]). Since most observational data sets lack measures of individual ability, the regression of earning on education has error which includes unobserved ability and hence is correlated with the regressor of education.

Based on this, presumed we can find κ_i variables that influence education directly, but on earnings indirectly only through the impact on education. We can then rule out this confounding variables bias. The κ_i variables are called instruments or instrumental variables, and this method is known as the instrumental variables method (IV method) which is a general approach to control for endogeneity. Several instruments have been used in the literature, for example, distance to educational institutions or family background variables. In particular, parental education is documented as one of the most commonly involved variable in this context. It is necessary to note that parental education could be used either as an instrument for education or as a proxy for the unobserved ability. Existing studies use it in both ways. For instance, when it is used as an instrument (see e.g., Ashenfelter and Zimmerman, 1997), it is assumed to be uncorrelated with the ability of their children but correlated with their education. Furthermore, economists have also tried to use a variety of 'nature experiment' to estimate the causal effect of education as it is not possible to set up an actual experiment as it is, for example, in the medical

[36] According to Cameron and Trivedi (2005), when variables are omitted from a regression, and when omitted factors are correlated with the included variables, a confounding bias results. Possibly confounding variables include not only ability, but also other variables such as family background characteristics (Maluccio, 1998).

literature. Such instruments constructed from the 'nature experiment' including, for example, individual's quarter of birth, which is uncorrelated with earnings except through their effect on education via school-start age policy and compulsory school attendance law (e.g. Angrist and Krueger, 1991).

However, the IV method can not always be applied to the empirical work as either necessary instruments may not be available, or they suffer from weak instruments bias. We thus need to be very cautious with respect to the validity and relevance of these instruments as applied in the empirical analysis. Instruments are valid provided they are orthogonal to the error term ($E(\kappa_i\varepsilon_i) = 0$). In the traditional simultaneous equations approach, this occurs if they are excluded from the equation of interest. Recently, the relevance of the instruments has received increased attention by researchers. It has been suggested that many applications of the IV regressions suffer from 'weak instruments' or 'weak identification', which is that the instruments are only weakly correlated with the included endogenous variables (Stock, *et al.*, 2002). If the instrument is extremely weak, it may provide little or no useful information (Imbens and Rosenbaum, 2005). The problem of weak instruments is quite prevalent in recent IV studies, and researchers have pointed out that some IV studies have been undermined by a lack of precision in their first stage estimates. Bound, *et al.* (1995) pointed out that even a weak correlation between the instrument and the error term in the earning equation can result in a large bias in the IV estimates even in large samples, and this problem is compounded if the instrument is weak. They argued that instrumental relevance is ensured only if there is a strong correlation between the instruments and the endogenous regressor. If this is not the case, IV estimates are asymptotically centered on least squares (Han and Schmidt, 2001).

According to Card (1994), an alternative to the use of instruments for education attainment is to assume that the pair of residuals ε_i and μ_i in equation (3.9) and (3.10) have a restrictive covariance structure, and then use either repeated observations over time for the same individual or observations for different individuals from the same family to 'difference out' the correlation between ε_i and μ_i. Estimates based on the twins fixed effects are mostly applied in this context, presumably that specifically monozygotic twins are genetically identical and have a similar family background, thus the effects of unobserved ability or family background should be similar for both twins (Li, *et al.*, 2005). They further argued that provided that at least some pairs of twins have different levels of education, a within-twin-pair estimator analogous to the standard fixed effects estimator removes the possible bias that is caused by the omitted ability and family background variables (e.g. Ashenfelter and Krueger, 1994). Nonetheless this still raises the question as to why, e.g., one twin goes to the university and the other does not. Is this distinction completely random or are they unobserved causal factors? If the latter, how do they affect the estimates?

The aforementioned arguments are focused on the possible solutions for ability bias. The main debate is whether the impact of education on earnings isolated the effects that are caused by education from the consequences of unmeasured, for instance, ability. In other words, whether the coefficient of education is biased in the sense that it is picking up both the impact of education and ability. We will return to this in detail in chapter 4.

3.1.4. Signaling/Screening[37]

The positive correlation between education and earnings has been explored with a variety of data sources and analytic techniques by a large amount of researchers in the literature. Economists offer two major theories explaining this correlation: human capital theory which we have discussed and signaling/screening hypothesis. Both of them assume that an individual obtains education up to the point where the marginal benefit given by the earnings function just equals the marginal cost of education. The theories differ in the reason that employers pay a premium for education (Kroch and Sjoblom, 1994). According to human capital theory, education is assumed to raise individual's productivity and subsequently future earnings. In this sense, the coefficient on education in the log earnings regression has been interpreted as an estimate of the productivity-enhancing effects of education. However, do workers' earnings differentials accurately reflect their different marginal productivities? There has developed the

[37] According to Weiss (1995), both signaling and screening serve to 'sort' workers according to their unobserved abilities based on information asymmetry. For example, students will choose a length of schooling to 'signal' their ability to employers, and the employers will demand a minimum level of schooling from applicants in order to 'screen' their workers.

signaling/screening hypothesis by Arrow (1973) and Spence (1973, 1974) that employers use educational attainment to identify individuals with certain valuable innate traits that cannot be observed directly: education is not fundamentally productive; nevertheless in equilibrium wages increase with education, since individuals of greater productivity experience less disutility from schooling and obtain more of it, thus signaling to employers their greater worth[38]. Simply speaking, education acts as a signal of individuals' innate traits which enable employers to identify workers. In the literature, Groot and Oosterbeek (1994) have tested these two competing theories by separating the number of actual years of education into five components: effective years, repeated years, skipped years, inefficient routing years and dropout years. Their results provide strong support for the human capital theory and refute the screening hypothesis. In addition, Kroch and Sjoblom (1994) used a similar method and obtained the same results.

It has been suggested in the literature that signaling/screening models extend human capital theory models by allowing for some productivity differences that firms do not observe to be correlated with the costs or benefits of education[39] (Weiss, 1995). Hence,

[38] For detailed discussions see Kroch and Sjoblom (1994). Game theoretic refinements for this have been suggested by Riley (1975), Mailath (1987) and Cho and Kreps (1987).

[39] Weiss (1995) argued that if the signaling/screening extensions of human capital theory are important, the marginal benefits from education for an individual could greatly exceed the expected value of the effects of those activities on productivity. For example, an accurate measure of the change in earnings for an individual who goes to school for 12 years instead of 11 would not measure the effect of that year of education on his productivity, but rather the combined effect of one additional year of learning and the effect of being identified as the type of person who has 12 rather than 11years of education.

the strong screening hypothesis (SSH) and the weak screening hypothesis (WSH) have been separately identified. The SSH presumes that education is being used exclusively as a signal immutable to productivity (Psacharopoulos, 1979). The WSH presumes that the primary role of education is to signal, but it may also augment inherent productivity (Arrow, 1973; Spence, 1973; Stiglitz, 1975). For example, Brown and Sessions (1998) found evidence in support of weak, but not strong, screening which indicates that education both signals and augments individual productivity. Psacharopoulos (1994) concluded, regardless of whether schooling signals or arguments productivity, it certainly enhances lifetime earnings and, as such, may represent a good investment for individuals.

The standard procedure to finessing this problem is to distinguish those situations where individuals would benefit from signaling ability from those situations where one would not expect such a benefit, and then to compare the rates of return to education of the two groupings (Kroch and Sjoblom, 1994). It has been documented in the literature that the possibly comparative groups include employees *vs.* the self-employed or public *vs.* private sector employees. Thus the difference between the returns to education for employees *vs.* the self-employed or between public *vs.* private sector employees is the value of education as a signal (Harmon, *et al*., 2003). In addition, the direct approach to distinguishing between ability and productivity is to include ability measures. It is noticeable that instead of the IQ scores as a proxy of ability as mentioned before, these ability measures need to specifically indicate the ability of making money, and they further need to be uncontaminated by the effects of education or they will pick up the productivity enhancing effects of education (Harmon, *et al*., 2003). It seems clear that these measures cannot be easily observed.

3.1.5. Experience[40] and the Schooling Vintage

According to Neuman and Weiss (1995), there are two types of human capital depreciation: *internal* depreciation and *external* depreciation. Internal depreciation which is assumed in most studies is the depreciation attributable to the worker including the loss of physical ability as well as any loss of mental capacity. External depreciation is caused by external forces, i.e., it is the lowering of the value of an individual in the market due to changes in the job or work environment. Analogously, De Grip and Van Loo (2002) considered that, from an economic point of view, two types of depreciation may be distinguished, *technical* skills depreciation that refers to a worker's loss of human capital, and *economic* skills depreciation that refers to the loss of market value of a worker's qualifications. Among the first type, they include the depreciation directly attributable to the worker's physical deterioration (*wear of skills*) and that due to *atrophy* of skills due to the lack of insufficient use of skills. As for the economic depreciation they further distinguish three causes, job specific skills obsolescence due to technological or

[40] Here we focus on schooling vintage on Mincer's (1974) experience-earning profiles. For detailed discussions see Neuman and Weiss (1995).

organizational developments, obsolescence related to shifts in the sector structure of employment and firm-specific skills obsolescence.

In particular, one element of these external/economic forces is the gradual obsolescence over time of the stock of knowledge obtained during the schooling process. This is due to the so-called ‘vintage effects’, which are closely associated with the pace of technological change (Ramirez, 2002). Simply speaking, the ‘human capital vintage effect’ means that the skills individuals have learned when they were young can become outdated and no longer productivity-enhancing as people become older (Skirbekk, 2002). Mincer (1974) changed the emphasis from age to labour market experience, although the depreciation rate was not identified in his typical model. Even so, researchers have attempted to estimate the depreciation of human capital under certain identification assumptions. Most studies which have dealt with it empirically have focused on how schooling vintage can affect Mincer’s experience-earning profile. A simple example would be if there is no vintage effect of education, returns from an extra year of recent schooling and an extra year of schooling in the distant past are the same. In addition, Neuman and Weiss (1995) argued that it seems unreasonable to assume that an elementary school graduate’s human capital suffers the same degree of obsolescence as a university graduate of electrical engineering. It is clear that the material taught in elementary schools has not changed much over time. However, an electrical engineer who was trained 20 years ago learned a lot of different material than one who just left the university. They introduced interacted variables[41] of schooling and experience with high-tech/low-tech dummies in the earnings equation and concluded that the rate of return to education slacks over time and also there is greater depreciation in the return to experience for the better educated. Thus the higher level of education, the more quickly the human capital becomes obsolete. Further more, this depreciation, driven by technical change, is more significant in industries strongly affected by technology growth (high-tech), such as the information technology (IT) and the electronics industries, although it could be argued that IT now impacts on most industries.

3.2. Health Capital

3.2.1. Grossman's Model (Grossman, 1972)

Investment in education is not the only form of human capital investment that individuals undertake. Health can be viewed as one form of human capital, and health capital differs from other forms of human capital. An individual's stock of knowledge affects his/her market and non-market productivity, while his/her stock of health determines the total amount of time he/she can spend producing earnings and commodities (Grossman, 1972).

[41] Instead of using interacted variables, an alternative way to take into account this possible vintage effect is to estimate the earnings equation for different groups classified by years of work experience to obtain the different rate of return to education over time. For example, Liu (1998) has implemented this approach and argued that the rate of return to education diminishes for more experienced individuals which indicated a higher return for more recently obtained education.

It is necessary to notice the distinction between 'commodities' and 'market goods'. Consumers produce commodities with inputs of market goods and their own time. In other words, consumers do not purchase health from the market but produce it, spending time upon health improving behavior or otherwise as well as purchasing medical inputs. The demand for medical care is a derived demand for an input to produce health.

Following the standard model of Grossman (1972, 2000), we assume that the utility function for a representative consumer is as follows:

$$U = U(\phi_t H_t, Z_t), \qquad t = 0, 1, ..., n \tag{3.11}$$

Where H_t is the stock of health at age t or in time period t, ϕ_t is benefit produced by one unit of health capital, $h_t = \phi_t H_t$ is the total health 'consumed' at age t, and Z_t is consumption of another commodity.

The initial stock of health capital H_0 is exogenous. H_t at any other age and the length of life n are endogenous. Length of life is determined by the quantities of health capital that maximize utility subject to production and resource constraints. The following equation describes the change of health capital:

$$H_{t+1} - H_t = I_t - \delta_t H_t \tag{3.12}$$

where I_t is the gross investment in health and δ_t is the rate of depreciation during the tth period. The rate of depreciation is exogenous but may depend on age. Consumers produce gross investment in health and the other commodities in the utility function according to a set of household production functions:

$$I_t = I_t(M_t, TH_t; E) \tag{3.13}$$

$$Z_t = Z(X_t, T_t; E) \tag{3.14}$$

In these equations M_t is a vector of goods purchased in the market, such as health care service, which is used to contribute to I_t. X_t is a similar vector of goods inputs that contribute to Z_t[42]. TH_t and T_t are time inputs, i.e. the amount of time devoted to the respective activities, and E is assumed to be other exogenous components of human capital exclusive of heath capital, such as education[43].

Furthermore, the consumer faces the following budget constraint:

---

[42] Hence we exclude the possibility that consumption of X_t can impact on health, as may be the case with excessive eating or drinking, or goods such as holidays which reduce stress.

[43] Grossman (2000) explained that the semicolon before E highlights the difference between this variable and the endogenous goods and time inputs. In addition, he assumed that E in equation (3.13) and (3.14) does not vary over the life cycle.

$$\sum_{t=0}^{n}\frac{P_t M_t + Q_t X_t}{(1+r)^t} = \sum_{t=0}^{n}\frac{W_t TW_t}{(1+r)^t} + A_0 \qquad (3.15)$$

where P_t and Q_t are the prices of M_t and X_t, W_t is the hourly wage rate, TW_t is hours of work, A_0 is initial assets, and r is the market rate of interest. The time constraint requires that Ω, the total amount of time available in any period, must be exhausted by all possible uses is:

$$TW_t + TH_t + T_t + TL_t = \Omega \qquad (3.16)$$

where TL_t is time lost from market and nonmarket activities due to illness and injury[44]. To sum up, equations (3.11) to (3.16) constitute the basic Grossman model and they jointly determine the demand for health.

Individuals inherit an initial stock of health that depreciates with age and can be increased by investment. This approach developed an important link between the household production theory of consumer behavior and the theory of investment in human capital: health is a choice variable because it is a source of utility (satisfaction) and also because it determines income or wealth levels[45] (Grossman, 2000). Gross investment in health

[44] Although we note in passing that this too could be argued to be dependent on other activities, both health and non-health related.

[45] Equations (3.11) to (3.16) do not include the 'full wealth' constraint. For detailed discussions see Grossman (2000).

capital is produced by household production functions. Hence the output of health is impacted upon by health-related choices including diet, exercise, smoking, etc. In addition, it is also affected by the factors which influence the efficiency of the production process for a given consumer as reflected by his or her personal characteristics, including factors such as age, gender, and possibly the level of education.

Health is both produced and demanded by consumers. Health is both an investment and a consumption good, and it relates to both a pure investment model and a pure consumption model in Grossman's monograph on the human capital model. In the former model the psychic rate of return is assumed to be zero, while in the latter the monetary rate of return is assumed to be zero. Empirical work mostly focuses on the estimation of the investment model based on its *powerful predictions from simple analysis* and a more adequate representation of people's behavior (Grossman, 2000). The stock of health and gross investment respond to variations in a number of exogenous variables[46]. Grossman examined two types of variations: evolutionary (differences across time for the same individual) and parametric (differences across consumers of the same age). Based on this, the wage rate, the price of medical care, the stock of human capital (mostly years of formal education completed) and age, and the unobserved depreciation rate in the initial period are included in the estimating equation. The investment model predicts that the coefficients of the wage rate and the stock of human capital are positive, while the

[46] For example, education is assumed to be exogenous here (see Grossman, 1972).

coefficients of the price of medical care and depreciation rate are negative[47]. In the empirical analysis, the stock of health is mostly measured by individuals' self-reported health, and the estimated equation mostly involves education level, age, the wage rate, etc. restricted by the data availability.

3.2.2. The Causality between Education and Health

An extensive review of literature conducted by Grossman and Kaestner (1997) suggested that years of formal education completed is the most important correlate of good health, whether health levels are measured by self-evaluation of health status, mortality rates or physiological indicators of health. In a broad sense, the observed correlation between education and health may be explained in one of three ways according to Grossman (2000): firstly, there is a causal relationship that runs from increases in education to increases in health; secondly, the direction of causality runs from better health to more education; thirdly, there is no causal linkage between education and health, instead, differences in one or more 'third variables' affect both education and health in the same direction. These explanations suggest that education clearly is an endogenous variable[48].

---

[47] For detailed discussions see Grossman (2000).

[48] Note that income is another endogenous variable in this context and the main direction of causality between income and health is also open to debate. Similarly, there are two main issues that need to be tackled in order to establish causality. These relate to: i) there are likely to be unobserved individual characteristics, which jointly determine both income and adult health; and ii) there exists the issue of reverse-causality, in which case people in good health are more likely to be economically productive and have higher incomes (Frijters, *et al*., 2005). We will review

We will start with the first one. More education leads to better health when more educated individuals are more efficient producers of health. This efficiency effect can take two forms: *productive efficiency* pertains to a situation in which the more educated obtain a larger health output from given amounts of inputs; *allocative efficiency* pertains to a situation in which education increases information about the true effects of the inputs on health (Grossman, 2000). For example, the more educated may have more knowledge about the harmful effects of cigarette smoking. In addition, education is also linked to income and occupation, two other components of socioeconomic status, and the gross effect of education on health may partly include these effects too[49].

It is clear that the second and third explanations challenged the causal relationship between education and health. According to Grossman (2000), the direction of causality may run from better health to more education because healthier students may be more efficient producers of additions to the stock of knowledge through formal schooling. Furthermore, this causal path may have long-lasting effects if past health feeds into current health status. Thus a positive relationship between health and education may reflect reverse causality in the absence of controls for past health, or more broadly, health endowments shaped prior to educational attainment. For example, Case, *et al.* (2003)

---

these issues in detail in chapter 5.

[49] However, as argued by Grossman and Kaestner (1997), the significant portion of the gross education effect on health can not be traced to the relationship between education and income or occupation.

confirmed that children who have experienced poorer health in childhood have significantly lower educational attainment and poorer health as adults. This reverse effect would create a positive simultaneity bias in measuring the effect of education on health (Groot and Maassen van de Brink, 2006).

The 'third variable' explanation is particular relevant if there exists a large unexplained variation in health after controlling for education and other determinants. This hypothesis has received a great deal of attention in the literature because it is related to the hypothesis that the positive effect of schooling on earnings is biased by the omission of ability. Fuchs (1982) identifies time preference as the third variable. Individuals' preferences for the future affect the likelihood of both educational attainment and good health behavior. Fuchs argued that persons who are more future-oriented attend school for longer periods of time and make a larger investment in their health[50]. Fuchs (1996) further argues that education is correlated with time preference, and that it is time preference that affects health rather than education. This hypothesis is tested by Sander (1995), who found the support for both the argument that education affects health and for the hypothesis that time preference matters. In addition, another important aspect that should be noted is the role of intermediate variables in the relation between education and health, such as parental education (Groot and Maassen van de Brink, 2006). Parental

[50] Fuchs also mentioned the possibility that schooling has an effect on time preference (preferences are endogenous), but he could not distinguish the direction of the causality.

education has an effect on health through its effect on educational achievements[51]. Grossman (2000) concluded that both the reverse causality explanation and the 'third variable' explanation indicate that the observed relationship between current health and education reflects an omitted variable which challenges the causal relationship between education and health. In econometric terminology, both explanations fall under the general rubric of biases due to unobserved heterogeneity among individuals (Grossman, 2000).

Several approaches can be employed to correct for these biases. Firstly, one can include past health measures in regressions that relate adult health to own schooling. The problem of this procedure, as argued by Grossman (2006), is that measures of past health may not be available or may be measured imprecisely. Secondly, unmeasured third variables can be controlled by examining differences in outcomes due to differences in education between siblings or twins. This approach is clearly restricted by the data availability. In addition, differencing between twins exacerbates biases due to measurement error (Neumark, 1999). Finally, one can employ the IV approach, which is the most commonly used procedure in the literature. As argued before, the IVs need to be correlated with education, but not be correlated with such omitted variables as ability, other inherited genetic traits, and time preference to obtain consistent estimates.

---

[51] For detailed discussions see Groot and Maassen van de Brink (2006).

3.2.3. Body Mass Index

The BMI is a practical measure used to determine overweight, developed by a Belgian statistician and anthropometrist, Quetelet in 1996. It is highly correlated with body fat, and, subsequently, health risk, specifically type 2 diabetes and cardiovascular disease, which are rapidly becoming major causes of death in adults in all populations (WHO Expert Consultation, 2004). The BMI is generally calculated as weight in kilograms over height in meters squared (weight(kg)/height(m)^2). The WHO has devised a classification where by persons with BMIs below the range 19-25 are considered underweight, those with BMIs above this range are considered overweight or 'at risk' and those with BMIs greater than or equal to 30 are considered obese. These WHO BMI classifications of overweight and obesity are intended for international use.

Recently, a growing body of literature in anthropometry indicates that these cut-off points are likely to be lower among Asian populations because the greater prevalence of cardiovascular disease risk factors is at lower BMIs in Asian populations than in Western populations (see, e.g. Moon, *et al.*, 2002). There is no clear single cut-off point for all Asians for high risk, and individual countries should set their own cut-off for high risk on the basis of their morbidity and mortality data (Choo, 2002). For instance, the overweight status among Chinese is defined as a BMI$\geqq$24 (Wildman, *et al.*, 2004). In addition, it has also been suggested that the BMI is age and gender dependent when used as an indicator of body fatness (see, e.g., Gallagher, *et al*., 1996; Burkhauser and Cawley; 2008). Thus the BMI, used indiscriminately for all people of all ages might provide misleading information. For example, when people are getting old, their body fat increases and muscle diminishes, while the BMI stays stable during these changes (Prentice and Jebb, 2001). However, there are no clear cut-off points after adjusting age and gender to be found in the literature.

3.3. Social Capital

3.3.1. The Dimensions and Measurement of Social Capital

Social capital, in short, refers to social connections and the attendant norms and trust (Putnam, 1995). The working definition of social capital that is emerging in an increasingly interdisciplinary literature refers to the networks, norms and understandings that facilitate cooperative activities within and among groups of individuals (Helliwell, 2001). Similarly, Halpern (2005) argued that most forms of social capital consist of a *network*, a cluster of *norms, values, and expectancies* that are shared by group members; and *sanctions* that help to maintain the norms and network. Accepting that there is no single form of social capital, as argued by Putnam (2001), we need to think about the multiple dimensions of social capital. Some forms of social capital are highly formal, like a labour union, or a national organization of any sort, formally organized with a chairman and a president, and membership dues and so on. However, some forms of social capital, such as a group of people who gather at bar, are highly informal. Both of these constitute networks in which there can easily be developed reciprocity, and in which there can be gains to the individuals in the networks (Putnam, 2001). Alternatively, some forms of social capital are very dense and interconnected, while some forms are very thin and almost invisible, although still providing networks and associated norms of reciprocity[52]. A particularly common and important distinction is between 'bonding' and 'bridging' social capital (Gittell and Vidal, 1998). The former focuses on ties within a group, sometimes in ways that increase distance and tension between groups; and the latter typically has a wider and more inclusive radius. According to Helliwell (2001), many actual situations are characterized by social capital that mixes these two types in different proportions.

The most compelling empirical evidence in support of the importance of social capital comes from the household- and community-level, and the best way to think of community-level social capital is as the set of social resources of a community that increases the welfare of that community, or in the language of economists, utility (Glaeser, 2001). There is an older definition of social capital that is individual-based. Following James (1904), it would be possible to define individual social capital as the set of social

[52] Putnam (2001) argued that if you nod to people in the hall, they are more likely to come to your aid if you should have a fit or have a heart attack.

attributes possessed by an individual – including charisma, contacts and linguistic skill – that increase the returns to that individual in his dealings with others. Researchers have believed that thinking about individual social capital is a prerequisite for thinking about the formation of community social capital, and that decisions to invest in social capital are made by individuals, not communities (Glaeser, 2001). Since individual social capital is very close to human capital, much of the theory on investment in human capital can be applied to social capital.

The theoretical discussion has outlined a number of different dimensions of social capital, and it is rather difficult to ideally measure them as is, for example, years of education for human capital. However, the social capital literature has primarily focused on two questions. First, there is a set of questions about types of non-professional organizations that an individual belongs to. Generally, organizational membership is aggregated up to a variable that captures the number of different types of organizations to which an individual belongs. Glaeser (2001) argued that it is probably the best measure of social capital investment on the individual level. Second, there is a survey question on trust that typically asks: "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" In some work, a measure of generalized trust in others is part of what we meant by social capital, while other researchers prefer to treat social trust, or interpersonal trust, as something that is generated and supported by the more variable sorts of social capital (Helliwell, 2001).

Furthermore, the concept of social capital is a central feature of civil society and trust as a key ingredient of civil society is vital for sustaining economic wellbeing (Putnam, *et al.*, 1995). Inability to commit oneself to or believe in anything that transcends one's own private interests leads to the weakening of commitment in family and community, and in these associations, people build relationships of trust and mutual aid (Persell, *et al.*, 2001).

The education-social connection relationship should probably be seen as the most important fact about the formation of social capital (Glaeser, 2001). In all studies of differences among individuals as to the extent of their connectedness and trust, education levels are uniformly the strongest and most consistent explanatory factors. There are many possible interpretations of this relationship. Firstly, if education reflected a greater orientation towards the future (i.e. lower time discount factor), then we should not be surprised that people with more human capital also invest more in social capital. Secondly, social capital may be reflecting relative prestige and it is very likely that interacting socially may be more pleasant for more prestigious individuals. Finally, according to Glaeser (2001), a significant part of education is in learning social skills. A great deal of time is spent in schools learning how to deal with peers in an educational setting: in younger grades, learning cooperation exists as a major part of education process; and in college, a great deal more time is spent on constructive social activies.

3.3.2. Social Capital and Well-being

There are measurable consequences to social capital. To get a full picture of the consequences of social capital for well-being, we need look at different levels of individuals, communities and nations and different measures of well-being. Putnam (2000) reviewed a large number of empirical studies which report a positive linkage between social capital and child welfare, crime, neighbourhood vitality, health and happiness. One of the most convincing pieces of evidence of the positive impact of social capital lies in the area of personal health based on the individual level. With respect to our empirical work in this context, we will focus on those effects of social capital that flow through health outcomes to well-being, and further extend to the direct impact of some social capital variables on subjective well-being.

Like human capital, social capital provides important benefits to individuals and societies. Social capital in various forms is hypothesized to affect an individual's health, and good health is included in various ways in all broader measures of well-being. One of the earliest papers, according to Helliwell (2001), and still one of the most convincing, is a mortality study by Berkman and Syme (1979) of almost 5,000 residents of Alameda County, California, selected randomly in 1965 and followed thereafter for nine years. Four types of social contact were assessed: marriage, contacts with close friends and relatives, church membership, and informal and formal group associations. For each of the four types, individuals who had greater ties had lower mortality rates over the nine-year period[53]. In addition, the results of numerous studies indicate that there exists a link between social connectedness on the one hand, and health and personal well-being on the other (after controlling for social, racial and demographic characteristics of individuals) (OECD, 2001). Putnam (2000) suggested two possible reasons for these links: i) social networks furnish tangible assistance and care which reduce psychic and physical stress, and ii) social capital might trigger a physiological mechanism stimulating individuals' immune systems to fight disease and buffer stress.

Halpern (2005) argued that the factors affecting positive, subjective well-being are largely a mirror image of those driving everyday psychopathology. In this respect it is no surprise to find that happiness or life satisfaction is strongly affected by personal relationships. The married are very substantially happier than the unmarried, and those living together are happier than those living alone. Thus marriage serves as a buffer against the hardships of life and it provides emotional support which produces positive states of well-being (Eddington and Shuman, 2005). In addition, measures of social connection also have significant effects on happiness, like being a member of some voluntary associations (Argyle, 1987; Helliwell, 2003). As mentioned before, these associations constitute networks in which there can easily be developed reciprocity.

[53] It is necessary to note that these relationships are not all large and robust, and they could be subject to alternative interpretations, such as confounding factors. For detailed discussions see Halpern (2005): *Social capital*, page 75.

Specifically, failure to trust people means an inability to make use of potential help which affects the efficiency of the individuals' welfare maximization problem. However, it is important to note that these are not free of the risk of reverse causation or of the possibility that both social engagement, trust and well-being are the results of other, unmeasured, interpersonal differences (Helliwell, 2003). For example, the married are happier, or are happy people more likely to marry and stay married than others? In addition, it is very likely that people with optimistic personalities are likely to report higher levels of well-being and a more positive view of social interaction; hence they are more likely to engage in social interactions and to derive pleasure from them. More optimistic people also tend to be more trusting and satisfied with their lives. Theoretically, the approaches dealing with this possible bias between social capital variables and happiness would be similar to the ones we have discussed in the previous sections on education and earnings or health. Practically, it is very difficult to either find a suitable proxy to measure personality, or obtain suitably strong instruments.

3.4. Conclusion

This chapter introduced the theoretical frameworks and empirical foundations related to our empirical analysis in the following three chapters. We started with the theory and the model of human capital, and further expanded this to cover both health capital and social capital. The main areas of concern lies with the impacts of education on earnings, health, and some social capital variables and thence on subjective well-being. Note that there have not yet been clear measurements of social capital as, e.g. years of education serves for human capital. We focus mainly on the possible difficulties why these associations are difficult to test causally.

There are two main issues that need to be tackled in order to establish causality. These relate to: i) there are likely to be unobserved individual characteristics, which jointly determine both regressors and outcomes; and ii) there exists the issue of reverse-causality. Broadly speaking, both explanations indicate that the observed relationship between regressors and outcomes reflects an omitted variable. In the earnings equation, we mainly focused on the unobserved individuals' ability; in the health equation, we discussed the omitted individuals' inherent health and time preference; and for subjective well-being, it is very likely that certain types of personality influences both social capital variables and subjective well-being. The most common way to deal with this is applying the instrumental variable approach, although in many applications it is difficult to find suitably strong instruments and empirical work relating to education has in particular been the subject of substantial criticism in terms of its reliance on weak instruments. Data availability problems relating to the issue of finding suitable instruments are further compounded for transitional economies, such as China. Therefore, there is a need to supplement the traditional research in these areas with that based on alternative approaches to dealing with endogeneity. Throughout the thesis, we suggest alternative approaches, each specific to the context in which it is used. Their value lies more in that they present extra evidence on the impact of education, rather than they unambiguously provide a solution to the endogeneity problem. Neither approaches may be beyond criticism, but if different approaches lead to similar conclusions the value of those conclusions are strengthened. We will discuss these issues in detail in the following chapters.

Chapter 4 Private Returns to Education in Urban China

4.1. Introduction

The private return to education is a topic to which economists have paid a great deal of attention since Schultz (1961), Becker (1964) and Mincer (1974) laid down the theoretical foundations. The benchmark model for the returns to education is the key relationship derived by Mincer (1974), which involves regression of the natural logarithm of earnings against educational attainment and working experience. A great deal of work based on this model has been done for many countries. A large number of studies from all parts of the world show that educational returns for an additional year of schooling are positive and range from 5% in developed countries to as high as 29% in developing countries (Psacharopoulos, 1994). An early study on this topic for China is by Byron and Manaloto (1990). Using a sample of eight hundred adult workers from the city of Nanjing in 1986, they reported the rate of return to education to be less than 3.7%. Yu and Hannum (1996) used the national sample of 1988 and concluded the rate of return to education is 2.9% for males and 5.0% for females. It therefore seems that the return rate in China is, or was, relatively low. This result is probably because either China’s long term planned allocation of labour resulted in a relatively equal distribution of income, or the combination of a small sample size and inadequate variation in the data used (Liu, 1998).

China's transition from a planned to a market economy began at the end of 1978. By October 1992 when the 14th Congress of the Chinese Communist Party formally authorized the concept of the "socialist market economy", the Chinese economy in terms of its structure and institutional condition had changed almost beyond recognition (Lin, *et al.*, 1996). Economic reforms have brought significant changes to China's labour market, including market determined wage rates, which may have new implications for the returns to education. The main aim of this chapter is to examine to what extent this is so, and in particular to what extent returns to education have been increasing during the 1990s and to what extent such returns depend upon factors such as gender, sector and region.

According to human capital theory, education is assumed to raise individual's productivity and subsequently future earnings. However, there has developed the signaling/screening hypothesis by Arrow (1973) and Spence (1973, 1974) that employers use educational attainment to identify individuals with certain valuable innate traits that cannot be observed directly. Thus whether workers' earnings differentials accurately reflect their different marginal productivities is open to more debate. More precisely, it is difficult to determine whether the primary role of education is to raise productivity or to signal. Indirectly it is one of the objectives of this chapter to shed light on these issues. In this chapter, we use the panel data of the China Health and Nutrition Survey (CHNS) for the years of 1989, 1993, 1997 and 2000, covering more than a decade of Chinese economic reform, to estimate the returns to education. Such estimates of the rate of return to education, according to Duraisamy (2000), are a useful guide for private and even public investment in education.

This chapter is organized as follows: Section 2 reviews the literature based on different data sources and analytic techniques on returns to education as mentioned in chapter 3. In section 3, we outline our econometric framework. In section 4, we describe the dataset and model specification. Specifically, Mincer's model is slightly extended to include education depreciation, and our hypothesis is discussed. Section 5 presents the empirical results and finally we conclude this chapter.

4.2. Literature Review

Education plays an important role in the modern labour market. In human capital theory, education is an investment of current resources in exchange for future returns. Private investment in education theoretically provides an individual with the higher lifetime income level, social status and personal freedom (Filer, *et al.*, 1996). A large amount of literature in many different countries and time periods have confirmed that better-educated individuals earn higher wages, experience less unemployment, and work in more prestigious occupations than their less-educated counterparts (Card, 1999). The benchmark model for the development of empirical estimation of the returns to education is the key relationship derived by Mincer (1974). Empirically, this involves regressing the natural logarithm of earnings (Y) against educational attainment (edu) and working experience (exp). This is expressed as follows:

$$Ln(Y) = \beta_0+\beta_1 edu+ \beta_2 exp+ \beta_3 exp^2+\varepsilon \qquad (4.1)$$

A large literature has grown up around the Mincer model. Two of the most comprehensive surveys are presented by Psacharopoulos (1985, 1994). They cover the results of estimating the returns to human capital for over 60 countries. Psacharopoulos and Patrinos (2002) further updated the results based on estimates of returns to education for 98 countries. This study confirmed the results that the highest returns are recoded for low and middle-income countries. Returns to education for Asia are at about the world average: 10%. As suggested by the Mincer model, the explanatory variables are normally the years or the level of schooling, experience and experience squared. Given perfectly competitive markets, this should represent a full model. Partially to allow for market imperfections a number of other variables, primarily dummy variables, tend to be included relating to gender, sector and region, etc.

The estimates are generally derived using standard least squares regression with the explanatory observed variables that are deemed relevant by the researcher in addition to education attainment and experience. This interpretation is accurate as long as there is no linkage between education and the error term in the earnings equation. Education is an endogenous choice variable in the underlying human capital theory as it depends on the choices of educational investment made by individuals. Thus ordinary least squares estimates of its impact on earnings do not necessarily measure the causal impact of an exogenous change in education. As already discussed in chapter 3, estimates of returns to education based on the analysis of using IV measures are mostly applied in the literature.

However, many applications of the IV regressions in the literature suffer from weak instrumental variable bias (Bound, *et al.*, 1995).

4.2.1. Estimates Using Standard Least Squares Regression

In this section we selectively review the research which has been done for some transitional economies by standard least squares regression. It has been suggested that data availability makes the issue of finding suitable instruments rather severe for transition economies. The first transition to market economy conditions emerged in rural China around 1978, and the focus of the reforms has gradually shifted to urban areas since 1984. Up to the end of the 1980s, the economic, political and social changes in the former communist and socialist countries of Europe and Soviet Union were even more dramatic (Verhoeven, *et al*., 2005). There are at least two competing hypotheses about the returns to education during transition from centrally planned to market economy: one hypothesis is that labour market reforms at the onset would entail immediate increases in returns; the other is that the skills hitherto acquired can not be easily transferred to a changed economic situation, thus one would expect to see a temporary decline in returns during transition. In addition, the gender difference of the returns to education has also been the subject of the analysis. Women generally receive higher returns to their schooling investments (Psacharopoulos, 1994). Flanagan (1998) studied the Czech labour market and reported that returns to a year of schooling increased from 4.3 percent in 1988 to 5.7 percent in 1996. For males, the change in returns to schooling was from 3.7 percent to 4.5 percent, while for females this change was from 5.1 percent to 7 percent. Chase

(1998) found that the returns to one year's of schooling for men increased from 2.4 percent in 1984 to 5.8 percent in 1993 in the Czech Republic. For Czech women, he finds that the returns are higher than for men and increased from 4.2% in 1984 to 7% in 1993. Campos and Jolliffe (2002) used the data for about 2.9 million wage earners from 1986 to 1998 in Hungary and concluded that returns to schooling increased from 6.4% in 1986 to 11.2% in 1998. The evidence for the Russia Federation is somewhat different. Cheidvasser and Benitez-Silva (2000) found that the returns to education in the Russia Federation are quite low, less than 5%, and there is almost no improvement in returns to education in the post-reform period of 1992-99. In contrast, some studies found evidence of reduced returns to education in the former Republic of East Germany between 1988 and 1991 (see e.g., Krueger and Pischke, 1995).

There is some work which specifically relates to China. The most widely used datasets in estimating returns to education in China are the two waves of the Chinese Household Income Project (CHIP), which collected household survey data in 1988 and 1995. Maurer-Fazio (1999) used the urban sample of the 1988 CHIP and the Chinese Labour Market Research Project (CLMRP) data and concluded that the overall returns to education were estimated to be 2.9% and 4.5% for male and female workers respectively. Similarly, Liu (1998) used the CHIP 1988 data and found the rate of return to education to be 3.6%. He also considered differences in the returns to education across the provinces and found that Guangdong province[54] had the highest returns among the provinces. With respect to changes in the labour market, Li (2003) using the CHIP 1995

[54] Other provinces include Liaoning, Shanxi, Jiangsu, Anhui, Henan, Hubei, Gansu and Yunnan. Guangdong is one of the five southeast coastal provinces where economic reforms started early. The others are: Jiangsu, Shandong, Zhejiang and Fujian.

reported the overall rate of return to be 4.3% for men and 6.9% for women, which is higher than the CHIP 1988 results (for example, compared to 2.9% and 4.5% respectively by Maurer-Fazio (1999)). He classified workers into three cohorts depending on when they started working prior to economic reform, up to 1979, in the early stage of urban reforms, 1980 to 1987, or during the advanced stage of urban reform, 1988 to 1995. The average annual rates of return to college education were 7.7% for the pre-1979 cohort, 14.1% for the 1980-1987 cohort, and 14.8% for the 1988-1995 cohort. In contrast to Liu (1998)'s results on regional differences, he found that the rates of return were higher in a less-developed province, Gansu, than in a higher-income province, Guangdong. In addition, Wu and Xie (2002), based on the survey of "Life Histories and Social Change in Contemporary China", examined earnings inequality and earnings returns to education in China among four types of workers (stayers, later entrants, market losers and early birds as characterized by their labour market history). Compared to workers staying in the state sector, early market entrants no longer enjoy advantages. The commonly observed higher earning returns to education in the market sector are only limited to recent market entrants.

Researchers have also searched for evidence of increasing returns to education over time. Zhao and Zhou (2002) used a sample of urban residents drawn from 20 cities through a stratified sample in 1993 and 1994. They found significant and substantial increases in returns to education over time and especially in the non-state sector where market mechanisms prevail. Knight and Song (2003) also found that returns to college education rose from 4.9% in 1988 to 15.0% in 1995. This is consistent with the gradual development of both labour market reforms and economic reforms. In addition, a more recent study is by Zhang, *et al.* (2005), who used the data from fourteen consecutive years of Urban Household Surveys (1988-2001), and concluded that there had been a dramatic and consistent increase in returns to education in urban China. In 1988, the rate of return to education was only 4.0%, by 1999, it had risen to 10.2%. Similarly, returns to education are consistently higher for females than for males. In addition, they also found that the returns to schooling were lower in richer provinces, i.e., Guangdong, Zhejiang and Liaoning, which is consistent with Li's (2003) results, although this difference has declined over time. We summarize this in Table 4-1.

Table 4-1 Summary of the Selected Literature for Returns to Education in China

Authors	**Samples**	**Returns to Education**
Maurer-Fazio (1999)	CHIP 1988	Rates of return were 2.9% for men and 4.5% for women.
Liu (1998)	CHIP 1988	The developed province, Guangdong, had the highest returns among the provinces.
Li (2003)	CHIP 1995	Rates of return were 4.3% for men and 6.9% for women; Rates of return were higher in a less-developed province, Gansu, than in a higher-income province, Guangdong.
Knight and Song (2003)	CHIP 1988 and 1995	There is a substantial increase in all the education coefficients between 1988 and 1995; Rates of return were lower for profit-making than for loss-making enterprises.
Zhao and Zhou (2002)	A stratified sample from 20 cities in 1993 and 1994	Rates of return increase over time; Rates of return are higher in non-state sectors.
Wu and Xie (2002)	Life Histories and Social Change in Contemporary China (1996)	Higher earning returns to education in the market sector are only limited to recent market entrants.
Zhang, *et al.* (2005)	Urban Household Surveys (1988-2001)	Rates of return increase from 1988 to 1999; Consistently higher for females than for males; Rates of return were lower in richer provinces.

Although the research suggests that the rate of return to education varies widely overtime and between countries, most studies find evidence of increasing returns to education during the transitional process. This further suggests that such returns depend on supply and demand for different amounts or types of labour, and perhaps that in many cases the market is not in equilibrium and indeed also there is the possibility of market imperfections. With respect to China the results suggest an increasing rate of return to education over time and higher returns for women than men.

4.2.2. Review of the 'Vintage' Effect

De Grip and Van Loo (2002) have reviewed the economic literature on the various causes of skills obsolescence and the ways in which skills obsolescence has been estimated. One element of skills obsolescence is the gradual obsolescence over time of the stock of knowledge obtained during the schooling process. This is due to the so-called 'vintage' effect as mentioned in chapter 3. Early human capital theory dealt with skills obsolescence due to technological change[55]. For example, Su (1973) simply defined skills obsolescence as the gap between what a worker already knows and what he needs to know for working with the new technology. Su (1973) further gave the condition for the optimal point in time at which a firm terminates the retraining of existing workers and starts hiring new workers. However, Holtmann (1972) made the assumption that the gap is smaller for workers with a higher quality of initial education, since a higher quality of initial education reduces the training costs later in a worker's career. Rosen (1975)

[55] Other causes of skills obsolescence see De Grip and Van Loo (2002).

considered the obsolescence of skills due to technological development to be a 'vintage' effect: *Obsolescence occurs because stocks of knowledge available to society change from time to time... Sometimes new knowledge proves received knowledge to be incorrect or at least less general than was supposed at an early time. Similarly, production innovations often render useless skills associated with prior methods. In both cases, capital losses are imposed on those embodying the earlier knowledge and skills.* Van Imhoff (1988) further modelled the negative effect of a greying labour force on economic growth by means of a vintage model which takes account of skills obsolescence due to technological developments.

An extensive empirical literature on Mincer's earnings equations has evolved on the returns from investments in human capital, and empirical evidence is consistent with the existence of a process of human capital depreciation. Carliner (1982) used the Mincerian equation to look at the human capital depreciation process when studying the evolution of the salaries of U.S. males close to retirement with the concentration on worker's physical deterioration. Neuman and Weiss (1995) analysed the depreciation of human capital for the Israeli case. They tried to identify depreciation due to the economic depreciation caused by external forces, i.e., technological or organizational changes, which affects the value of the worker in the market. They extended the Mincerian model to include depreciation specifically linked to education or the 'vintage' effect. Since the 'vintage' effect is not the same in all sectors, while technical depreciation is, it is possible to identify both of them by comparing data from high- and low-tech sectors. Based on this, Ramirez (2002) carried out a similar study for Switzerland. A somewhat different result was obtained by Arrazola, *et al.* (2005), who concluded that for Spain, the depreciation rate does not vary according to the level of education, however, it does so depending on employment spells endured by the workers in recent years and on whether they have undertaken training courses.

4.2.3. Returns to Educational Signals

Accepting the positive relationship between education and future earnings, economists offer two major theories explaining this correlation: human capital theory and signaling/screening hypothesis. Both of them have been discussed in chapter three. Signaling and screening theories appear in the economics literature in the early 1970s as economists started to explore the consequences of imperfect information (Arrow, 1973; Spence, 1973). For example, when employers have imperfect information about the productivity of potential employees, one readily available piece of information is years of educational attainment. Therefore, in the absence of perfect (better) information, employers use these educational signals to predict the true productivity of workers in the hiring process. Because both of the theories explained the positive relationship between education and earnings as resulting from a positive relationship between education and productive ability, it is very difficult to empirically separate these two hypotheses. There have been a great many studies attempting to do so, using a wide variety of approaches. For the relevance of this chapter, we review a limited number of papers focused on the time path of the returns to educational signals to test for the existence of signaling.

Researchers have contested whether or not returns to education signals should decline as workers gain increased labour market experience. If employers learn about the true productivity of their employees over time, therefore causing the educational signal itself to be of less value, returns of education signals decline overtime. For example, Layard and Psacharopoulos (1974) argued that returns to educational signals will fall with experience as employers come to have better information about their employees' real productivity. Their examination of relevant studies of time showed no such decline, thereby proving evidence against the signaling hypothesis. However, Riley (1979) pointed out that a basic condition for a signaling equilibrium is that employers find out that their predictions based on the education signals are correct on average: some members of a particular educational signaling group will earn a higher wage and some will earn a lower wage, the return to an educational signal is constant over time. He uses this argument to dispute the hypothesis that returns to educational signals decline over time. Farber and Gibbons (1996) presented a model to incorporate Riley's point. They argued that as long as the new information acquired with work experience is orthogonal to the educational signals, the observed return to those signals will not decline. Belman and Heywood (1997) argued that the productivity of a worker in a particular job depends on the quality of the match between the requirements of the job and the abilities of the worker. In this sense, the worker's productivity is constrained by the quality of the job match. Employers can not immediately observe the true productivity of workers. The authors assumed that the matching process results in non-optimal matches in the first period, but the true productivity of the new employees is revealed in the second (and last) period. They calculated the return to an educational signal in each period to demonstrate that as workers gain experience the returns to such signals do attenuate over time. Furthermore, a recent paper by Habermalz (2003), who assumed that information is imperfect but symmetric in the first period, from period two onward workers become privately aware of their true productivity and the employer is still excluded from the knowledge of the true productivity of their workforce, offered a different explanation related to employer-learning from Belman and Heywood's (1997). His results indicated that for above median ability workers returns to educational signals first increase and then decrease, and empirically the return to an educational signal increases for a large portion of an individual's work life.

4.2.4. Estimates Using IV Measures and Weak Instrumental Variable Bias

Over the years there has been considerable interest focusing on the causality debate between education and earnings. A comprehensive survey is presented by Card (1999), which covers different techniques used in the estimation of the returns to education with the focus on a possible endogeneity bias in standard least squares estimates. The disturbance term ε in equation (4.1) captures unobservable individual effects and these individual factors may also influence the education decision. A common example is unobserved ability. More broadly, if there is a correlation between schooling and the error term in the Mincerian specification, schooling is endogenous and the estimated rate of return by standard least squares method will yield biased results. There have been a number of approaches dealing with this problem in the literature. The most commonly used method is the instrumental variables approach, which typically looks to the educational attainment literature for appropriate instruments or exploits some ‘natural experiment’ (Maluccio, 1998). Researchers assume the existence of instrumental variables uncorrelated with the structural error but correlated with the endogenous regressor. According to Imbens and Rosenbaum (2005), even in the best of situations, the assumptions that are required for an instrument are plausible, not certain. Specifically, if the instruments are only weakly correlated with the endogenous regressor then the IV estimates are potentially as biased as the standard least squares estimates.

One important example of the use of IV methods to make a causal inference is the paper by Angrist and Krueger (1991), which used quarters-of-birth as an instrument to circumvent bias in estimating the returns to education. They argued that quarters-of-birth are exogenous sources of variation in educational attainment due to existing compulsory schooling laws. Two findings support this association between quarter-of-birth and both age at school entry and educational attainment (Cruz and Moreira, 2005): First, the effect of birth on school attainment varies across states, depending on the legal drop-out age. The drop-out rate is considerably higher in states with an age-16 requirement compared with states with a longer requirement. Second, the relationship is weaker for more recent cohorts. Since the average level of education has increased over time, the laws are less likely to influence more recent cohorts. To estimate returns to education using quarters-of-birth as instruments, Angrist and Krueger (1991) used data from the public use data files, describing samples of individuals, from US censuses of 1960, 1970 and 1980, with sample sizes that were greater than 200, 000 for men born between 1920 and

1929, greater than 300, 000 for men born between 1930 and 1939 and greater than 400, 000 for men born between 1940 and 1949. However, the relevant sample size varies somewhat depending on the details of the different analyses that they performed (Imbens and Rosenbaum, 2005). They estimated earning equations considering specifications that combine different sets of instruments (quarter-of-birth possibly interacting with year-of-birth and region-of-birth) and different sets of covariates (race, metropolitan area, marital status, age and age squared). Their estimates for returns to education are sensitive to the different specifications but with remarkably small standard errors. For example, using the 1980 US census, the two-stage least squares (TSLS) estimates range from 6.0 percent to 9.9 percent for the 1930-1939 cohort with a standard error around 2 percent.

Bound, *et al.* (1995) replicated Angrist and Krueger's results for 1930-1939 cohort and questioned the reliability of quarter-of-birth as an instrument. Their criticism is two-fold. First, quarter-of-birth can be correlated with unobserved characteristics of the individual. Therefore, these instruments are not truly exogenous, being correlated with earnings after controlling for education. The second criticism is that quarter-of-birth is only weakly correlated with educational attainment in some specifications. Specifically, using exactly the same data, they replaced the quarter-of-birth information by entirely irrelevant random numbers, which should be useless for estimating the return to education, repeating this process 500 times, applying TSLS each time, obtaining a mean estimated return of a 6 percent increase in weekly wages for a year of additional education, with an estimated standard error around 2 percent as well. Moreover, by varying the instruments, they found greater apparent precision with more instruments than with fewer instruments, even though the instruments are all just irrelevant random noise (Imbens and Rosenbaum, 2005). Clearly, if the instruments are only weakly correlated with the explanatory variable, standard asymptotic theory can not be employed to developed reliable inference methods (Cruz and Moreira, 2005). The problem of weak instruments is quite prevalent in recent IV studies, and in most of the studies surveyed in Card (2000) the IV estimates suffer from imprecision and are not significantly different from the standard least squares regression.

4.3. The Data and Model Specification

4.3.1. Data Descriptions

The panel data we use comes from the Household Survey in China Health and Nutrition Survey, which is collected mainly by the Carolina Population Centre. It covers the four years of 1989, 1993, 1997 and 2000, and thus covers much of the period of China's economic reform process. This dataset is designed to examine the impact of the health, nutrition, and family planning policies and programs implemented by national and local governments. It contains information on wages, education, and other demographic information, so it is suitable to estimate the returns to education in China. We choose the urban area only, because of a large number of education missing values or no education attained in some/most rural areas. We restrict our sample to wage employees, currently working with only one primary job. The analysis was also restricted to those between 15 and 65. After also excluding observations with less than full information, this provided

5589 data points in the four years altogether. All the variables are defined in Table 4-2 and Table 4-3 shows summary information.

Table 4-2 Data Description of the Main Variables

Variables		**Data Description**
Ln(*Y*)		The natural logarithm of average monthly wages (inflation adjusted) which consist of basic wages, subsidies and bonuses.
Education (edu)		Formal education years (0-18)
Potential Experience (exp)		Experience=age–education–6
Gender		Male=1 Female=0
Sectors		Public: governmental units and state-owned enterprises/institutes; Private: individual/private enterprises and three-capital enterprises (owned by foreigners, overseas Chinese and joint ventures), and the remaining sector includes collectives and others.
Regions	1. Provinces	Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou
	2. Distance	Distance between the capital of upper nine provinces to Beijing and Shanghai respectively.

Almost 60 percent of all workers are employed in the state sector. This ratio increased from 1989 to 1993, and declined thereafter. This shrinkage of state sector employment coincided with the restructuring and privatization of SOEs. Compared to the state sector, the percentage of workers in the private sector increased significantly from 6.55% in 1993 to 17.03% in 2000, after Deng Xiaoping's tour to the South in 1992 promoting openness and reform. The private sector here includes individual/private enterprises and three-capital enterprises (owned by foreigners, overseas Chinese and joint ventures). The selection of these two sectors is mainly based on the rationale of the aggressive competition within the state sector and the emergence and development of the private sector as a prominent player in China's labour market since the early 1990s. People not in either the state or private sectors include, e.g., those in collectives.

Figure 4-1 Survey Regions

Source: China Health and Nutrition Survey. The emphasised (green) parts are the survey regions.

The survey covers nine provinces shown in Figure 4-1 that vary substantially in geography and economic development. These are Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou. Jiangsu and Shandong belong to coastal provinces, where economic reform started early; Liaoning and Heilongjiang belong to north-eastern provinces, and have many similarities in economic structure as China's major bases of heavy industry. Henan, Hubei and Hunan are middle provinces and Guangxi and Guizhou are western provinces where economic development is relatively slow. In 1989 and 1993, the survey includes eight provinces: Liaoning Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou. In 1997, Liaoning is substituted for Heilongjiang, and in 2000, the survey covers all nine provinces. In the regression results, there is one regional dummy variable for each region respectively. In addition, two metropolises have been chosen. One is Beijing, the capital of China, the other is Shanghai, the biggest commercial centre in China. The distance from the capital of the nine provinces to these two metropolises is included as well for reasons discussed later.

The dependent variable in our analysis is the natural logarithm of average monthly wages[56] which consist of basic wages, subsidies and bonuses. Subsidies in 1989 included those for food (grocery/meat), one-child, health, bathing/haircut and books/newspapers. In 1993, 1997 and 2000, the one-child subsidy was replaced by an accommodation subsidy. The others remained the same. Recall from the chapter 2, basic wages (or fixed income) in China are largely determined by the administrative system and reflect the principles of human resource allocation and rewards on a more or less rational basis (those with more seniority and education earn more) (Dickson and Rublee, 2000; Zhao and Zhou, 2002). In contrast, bonuses and subsidies, which constitute an increasingly large proportion of individuals' total income, are distributed less systematically. To some extent, bonuses and subsidies are dependent on the profitability of the enterprise itself. In the conventional context, the inclusion of bonuses and subsidies may be less closely

[56] In labour economics, the wage rate or weekly wage is often used because these wage rates measure more precisely productivity-related rewards. In the Chinese context, since most urban employees are full-time workers, the monthly wage serve the same purpose well (Zhao and Zhou, 2002).

associated with education, however, these employer-provided subsidies are more important in relatively high wage jobs, and their exclusion may affect the estimated returns to education, specifically for developing countries (see, e.g. Gibson and Fatai, 2006). In addition, we note that non-wage benefits such as a company car may also have an impact in developed countries. We use the consumer price indices from the World Development Indicators (2003) to deflate wage income in different years.

From the Table 4-3, we can see that the wages increased significantly from 1989 to 2000, and the substantial improvement is from 1997 to 2000. Table 4-4 and Table 4-5 show the specific distribution of wages across genders, sectors and different regions. Within the sample, there is a slight decline in the percentage of females in the dataset, from 45.32% in 1993 to 42.52% in 2000, and females earn less than males in all four years. In addition, individuals in state sectors earn less than those in private sectors except for 2000. It is necessary to mention that the sample size of the private sector in 1989 and 1993 is comparatively small. The state sector in the survey includes governmental units and state-owned enterprises/institutes. Regional inequality has always been an issue for China as argued in Chapter 2. Table 4-4 shows that individuals in coastal and middle provinces almost always have higher wages in all four years. People in Heilongjiang earn comparatively more as well, however, the selected sample in Heilongjiang includes only two recent years of 1997 and 2000. In addition, the distance between the capital of Jiangsu/Shandong and Shanghai/Beijing is the shortest of the nine provinces.

Table 4-3 Sample Size and Distribution

Variables	1989	1993	1997	2000	Full Sample
Female (%)	44.98	45.32	43.63	42.52	44.32
Average Age	37.79	39.00	37.63	39.05	38.32
Average Experience	23.77	23.86	21.84	22.62	23.17
Average Formal Education Years	8.02	9.14	9.79	10.45	9.16
Average Wages (RMB), Inflation Adjusted	312.09	405.29	544.04	841.20	486.56
State (%)	57.94	63.77	56.88	52.87	58.31
Private (%)	8.53	6.55	11.04	17.03	10.04
Married† (%)	76.03	70.46	79.50	70.13	74.17
Per cent Working†	65.95	66.46	57.63	55.51	61.98
Average Number of Household Members†	4.40	4.14	4.04	3.90	4.15
Liaoning (%)	16.79	19.12	-	8.28	12.26
Heilongjiang (%)	-	-	10.10	11.10	4.24
Shandong (%)	11.83	9.56	9.93	7.90	10.07
Jiangsu (%)	13.39	13.64	23.26	22.01	17.18
Henan (%)	13.55	10.76	12.65	11.38	12.20
Hubei (%)	11.23	12.83	14.52	11.67	12.44
Hunan (%)	12.04	13.24	13.58	12.61	12.79
Guangxi (%)	11.83	10.16	8.23	7.43	9.78
Guizhou (%)	9.34	10.70	7.72	7.62	9.04
No. of the observations	1852	1496	1178	1063	5589

Source: China health and Nutrition Survey, various years.
Notes: † No. of the observations for these three variables are slightly less than others due to missing values. They are only used for the test of possible selection bias in the empirical section.

Table 4-4 Sample Size and Distribution by Regions.

Province	1989 Average Formal Education Years	1989 Average Wages† (RMB)	1993 Average Formal Education Years	1993 Average Wages (RMB)	1997 Average Formal Education Years	1997 Average Wages (RMB)	2000 Average Formal Education Years	2000 Average Wages (RMB)
Liaoning	10.13	254.06	10.94	331.22	-	-	12.03	663.76
Heilongjiang	-	-	-	-	11.15	601.82	12.13	1058.64
Jiangsu	6.78	262.84	7.42	482.62	8.80	606.03	9.53	822.01
Shandong	6.95	405.85	8.46	589.38	9.34	449.99	10.20	1058.70
Henan	8.52	366.06	10.14	303.16	11.27	513.09	11.18	1025.99
Hubei	8.01	242.86	9.06	416.26	9.52	513.15	10.16	680.42
Hunan	8.64	370.58	9.51	473.07	10.43	672.26	10.53	868.67
Guangxi	6.63	364.64	7.84	372.81	9.08	455.56	8.77	519.80
Guizhou	7.66	231.35	8.59	311.18	9.27	380.37	10.11	785.22

Province	Full Sample Average Formal Education Years	Full Sample Average Wages (RMB)	Distance to Beijing‡ (miles)	Distance to Shanghai (miles)	No. of the Obs.
Liaoning	10.71	338.91	390	747	685
Heilongjiang	11.64	829.27	656	1051	237
Jiangsu	8.16	543.79	560	166	960
Shandong	8.32	559.05	228	459	563
Henan	9.97	500.42	395	517	682
Hubei	9.06	435.33	657	423	695
Hunan	9.63	559.82	834	543	715
Guangxi	7.71	405.44	1273	986	547
Guizhou	8.64	372.34	1000	939	505

Source: China Health and Nutrition Survey.
Notes: † Inflation adjusted; ‡ Calculations are based on the Meridian World Data system (http://www.meridianworlddata.com).

Table 4-5 Sample Size and Distribution by Sectors and Genders.

	1989		1993		1997		2000		Full Sample		
	Average Formal Education Years	Average Wages† (RMB)	Average Formal Education Years	Average Wages (RMB)	Average Formal Education Years	Average Wages (RMB)	Average Formal Education Years	Average Wages (RMB)	Average Formal Education Years	Average Wages (RMB)	No. of the Obs.
State	8.85	338.53	9.91	367.55	10.7	543.52	11.57	904.47	10.01	486.76	3259
Private	6.66	353.38	8.26	870.93	9.15	775.48	9.47	805.6	8.41	683.97	567
Female	7.63	278.1	8.8	359.54	9.58	462.75	10.27	752.7	8.84	425.31	2477
Male	8.35	339.88	9.42	443.22	9.95	606.97	10.59	906.67	9.41	535.31	3112

Source: China Health and Nutrition Survey.
Notes: † Inflation adjusted.

The independent variables include the number of years in formal education (edu), potential experience (exp) and experience squared. In the survey, completed years of formal education were measured by primary school (1-6 years), lower middle school (1-3 years), upper middle school (1-3 years), middle technical or vocational school (1-2 years) and college/university (1-6 years or more). For comparative purpose, we aggregate these discrete values based on China's education system to obtain continuous values of the formal education years[57]. From Table 4-3, we can see that average formal education years increased steadily from 8.02 in 1989 to 10.45 in 2000. More specifically as shown in Table 4-5, females have less formal years of education than males, although the values are very close in 1997 and 2000. Individuals in state sectors have substantially more years of education than those in private sectors in all four years. Potential experience is defined as: age-education-6, where 6 is the normal starting school age. It reflects the maximum number of years spent in work. Other explanatory variables in the earnings equation include gender, sectors, regional variables, and some interactive variables: gender*edu and sector*edu.

4.3.2. The Connection between Education and Earnings

As our baseline model, we begin by estimating returns to education with respect to wages income (Y) as follows:

---

[57] In China, basic formal education includes primary education (normally six years) and secondary education. Secondary education is divided into academic secondary education (normally three years of lower and three years of upper middle school) and specialized/ vocational/technical secondary education, i.e., after graduated from the lower middle school, one can apply for upper middle school or middle technical/vocational school.

$$Ln(Y) = \beta_0+\beta_1 edu+\beta_2 exp+\beta_3 exp^2+\beta_4 gender+\beta_5 region+ \varepsilon \quad (4.2)$$

Where Y is average monthly wages including basic wages, subsidies and bonuses. The inclusion of education, experience and experience squared is consistent with the standard human capital model – Mincer's earnings function. Coefficients β_0 to β_3 relate to this basic model. ε is a random error term. β_4 represents the basic earning differences between males and females. To an extent, as we have already suggested, these other variables are included to reflect labour and product market imperfections. For example, given perfect labour mobility the regional variables should not be significant. We use β_5 to capture the regional differences, and the variable region takes two forms: firstly a vector of regional dummy variables to represent the nine provinces; and the second form is two variables reflecting distance from Beijing and Shanghai respectively. We expand on the rationale for these later.

To further estimate returns to education, we then extend the equation (4.2) to incorporate information on sectors. It is necessary to include the sectoral differences in the Chinese context for reasons we have described in chapter 2. We also include two interaction terms: gender*education (genedu) and sector*education. We use the former to capture differences in the returns to education between genders, and the latter to identify and assess the specific channels that are effective in determining returns to education.

A major preoccupation in this context has been the potential for the unobserved determinants of education to be correlated with earnings. As mentioned in chapter 3, a large literature has developed looking at the causal impact of education on earnings. The main debate is whether the estimated impact of education on earnings has isolated the effects that are caused by education from the consequences of unmeasured, for instance, ability. The IV method is mostly applied in the empirical work to correct this bias presuming the valid and relevant instruments can be obtained. If this is not the case, for example, the instruments are only weakly correlated with the endogenous regressor, then the IV estimates are potentially as biased as the standard least squares estimates. Thus this approach is handicapped by the lack of suitably strong instrumental variables. This applies equally to other peoples' research and the database we are using. Instead we approach the problem of endogeneity from a slightly different perspective.

When labour economists estimate earnings equations, they almost universally find a significant positive partial correlation between years of education and earnings, and this finding is consistent with both the human capital and the signalling theories (Kroch and Sjoblom, 1994). We assume that the impact of education itself on wages may be because education either (i) *enhances* ability and productivity[58] or (ii) because it *reflects* ability. The former is based on the typical human capital theory that education adds directly to a person's productivity and that innate ability is merely a foundation to build on. The latter may be the case if education serves as a signal of innate productivity to alleviate the informational asymmetry between the employer and the employee about the true productivity of a worker (Habermalz, 2003). Therefore, the amount of education an

[58] There exists the possibility that education both augments and signals individual productivity based on weak screening hypothesis as mentioned in chapter 3 (see e.g., Brown and Sessions, 1998). Thus our hypothesis (i) indicates the primary role of education is to augment marginal productivity.

individual acquires signals the individual's innate ability. Considering cross-sectional data, we can summarize this within the context of the model by first arguing that the basic theory underlying the earnings equation is that wages are a function of ability[59] (A_i):

$$Y_i = f(A_i) \tag{4.3}$$

Ability itself is a function of basic or innate ability (A_i*) and education augmented ability:

$$A_i = A_i^* + h(E_i) \tag{4.4}$$

Included in E could be experience, although we will assume for simplicity that this is not so and restrict it to education. We also assume for simplicity linearity and separability. Finally education is a function of basic ability

$$E_i = g(A_i^*) \tag{4.5}$$

We cannot observe innate ability directly, although conceptually it is possibly to think that it could perhaps be measured. Hence when we estimate wages as a function of education we get:

---

[59] Human capital models do not naturally generate a positive correlation between ability differences and education (Weiss, 1971), however, the available evidence suggests that the benefit to schooling is greater for more able individuals (Weiss, 1995). We would expect this has the same implication for experience as well.

$$Y_i = f(g^{-1}(E_i) + h(E_i)) \qquad (4.6)$$

That is when we have no direct measure of innate ability and simply include education on the right hand side it is picking up the influence of both innate ability ($g^{-1}(.)$) and the impact of education in enhancing ability ($h(.)$). Unless one can find a suitable proxy for A* it is difficult if not impossible to separate out these two effects. Researchers have used the IV method or twin studies to isolate the causal effect of education on earnings. In reality this has not been totally satisfactory as argued before. However there are different implications in the model with respect to whether or not the impact of education declines with the period since the individual left formal education. If we assume potentially the only impact of education is to proxy innate ability ($g^{-1}(E_i)$) and if this is stable and unchanging over time, then the impact of education will not change as the individual ages[60]. If, however, education enhances ability, in part because of its impact on skills, then in a changing world as these skills date the impact of education on earnings should also decline. This is also consistent with the effect of schooling vintage. Thus, if we include an experience-education interaction variable in equation (4.2), a test for the impact of education on ability rather than simply proxying ability is provided by the existence of a significant negative sign on this variable. We will test this by panel data techniques which we are going to discuss in the next following section. It is important to emphasize that this rests on the view that the value of education, as a signal, does not decline over time. We reviewed the literature on this earlier, although not entirely unanimous, there is a tendency to accept this assumption. As Riley (1979) and Farber and

[60] We reviewed the literature of returns to educational signals in section 4.2, here we incorporate the assumption that the return to an educational signal is constant over time (see Riley (1979) and Farber and Gibbons (1996)).

Gibbons (1996) argued that a basic condition for a signaling equilibrium is that employers find out that their predictions based on the education signals are correct on average: the return to an educational signal on average is constant over time.

We acknowledge the possible endogeneity of the education attainment in the sense of unobserved ability in the OLS regressions and the bias it can cause. However, to some extent we question the hypothesis of endogeneity in the simultaneous sense. Education predates earnings and hence it is difficult to argue that they are determined by earnings per se. Yet there may be a linkage which goes beyond that of education enhancing skills, which further enhance earnings, but this will be because as we have argued, education may proxy ability. The special circumstances of China at this time further weaken the reverse link between education and earnings in that all the evidence is that prior to the 1990s, the earnings or financial return to education was much more limited than in a market economy. Hence that part of ability which is reflected in motivation to earn high incomes would have been less influential in impacting on people's human capital decisions prior to 1990 and these people constitute the bulk of our data base in all four years.

4.4. Econometric Framework - Panel Data Analysis

Data sets that combine time series and cross sections are called panel or longtitudinal data sets. Panel data provide information on individual behaviour both across time and across individuals. A major advantage of panel data is increased precision in estimation, more efficiency and less multicollinearity (Cameron and Trivedi, 2005; Baltagi, 2005). Heterogeneity across units is central to the issue of analysing panel data. Baltagi (2005) summarises the core issues in terms of econometric methodology, and Greene (2003) demonstrates more accessible illustrations of its applications. The following discussion draws heavily on Greene (2003)'s analysis within the context of our empirical work in the next section. The basic framework is a regression of the form:

$$Y_{it} = X_{it}\beta + Z_i\pi + \varepsilon_{it} \tag{4.7}$$

X_{it} has k columns and does not include a constant term. The heterogeneity or individual effect is $Z_i\pi$ where Z_i contains a constant term and a set of individual specific variables, which may be observed or unobserved, all of which are taken to be constant over time t. We apply our regression equation to the panel data model:

$$Ln(Y_{it}) = \beta X_{it} + Z_i\pi + \varepsilon_{it} \tag{4.8}$$

for i=1, 2, …,n (n=2891) and for each i, t=1, 2, …,T (T<=4[61]). The dependent variable is the natural logarithm of average monthly wages. β includes explanatory variables specified in equation (4.2). The residual ε_{it} has the normal properties. $Z_i\pi$ denotes the unobservable individual specific effects. We consider two cases:

[61] Unbalanced panel, we will explain this later in this section.

Fixed Effects (FE): If Z_i is unobserved, but correlated with X_{it}, then the least squares estimator of β is biased and inconsistent as a consequence of an omitted variable. However, in this instance:

$$Ln(Y_{it}) = \beta X_{it} + \alpha_i + \varepsilon_{it} \tag{4.9}$$

Where $\alpha_i = Z_i\pi$, embodies all the observable effects, specifies an estimable equation. This takes α_i to be an individual-specific constant term in the regression model. It should be noted that the term 'fixed' as used here indicates that the term does not vary over time.

Random Effects (RE): If the unobserved heterogeneity, Z_i, however formulated, can be assumed to be uncorrelated with X_{it}, then:

$$\begin{aligned} Ln(Y_{it}) &= \beta X_{it} + \mathrm{E}[Z_i\pi] + \{Z_i\pi - \mathrm{E}[Z_i\pi]\} + \varepsilon_{it} \\ &= \beta X_{it} + \alpha + u_i + \varepsilon_{it} \qquad (4.10) \end{aligned}$$

This random effects approach specifies that u_i is a group specific random element which although random between groups is constant for that group throughout the time period.

4.4.1. Fixed Effects

Fixed effects model assumes that differences across units of observation can be captured in differences in α_i as shown in equation (4.9). Therefore, α_i is the vector of individual dummies[62]. Note that when n is very large as in our case, regressions including (n-1) dummies may not be feasible, and taking deviations from individual-specific averages is equivalent to taking residuals from regression of $Ln(Y_{it})$ and X_{it} on individual dummies and then working with the residuals. Thus taking the average overtime yields:

$$\overline{Ln(Y_i)} = \beta \overline{X}_i + \alpha_i + \bar{\varepsilon}_i \tag{4.11}$$

Subtracting this from $Ln(Y_{it})$ in equation (4.9) yields the within model:

$$Ln(Y_{it}) - \overline{Ln(Y_i)} = \beta(X_{it} - \overline{X}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \tag{4.12}$$

for i=1, 2, …,n (n=2891) and for each i, t=1, 2, …,T (T<=4). Equation (4.9), (4.11) and (4.12) provide the basis for estimating β. Estimation (4.12) is known as the within groups estimator also known as the fixed effects estimator and it is performed by OLS. This is a classical regression model and no new methodology is needed to analyse it. If we want to investigate the differences across individuals, then we can test the hypothesis whether these differences are all equal with an F-test. The null hypothesis is: $\alpha_1 = \alpha_2 = ... = \alpha_{i-1} = 0$ against the alternative hypothesis that all unobserved individual-specific effects exist. If the F-statistic rejects the null hypothesis, the F-test indicates that unobserved individual-specific effects are different, and then the fixed effects model is more appropriate than pooled OLS.

[62] The model has been estimated with an overall constant and (n-1) dummy variables and each dummy variable coefficient will now be an estimate of difference with constant term where dummy "1" is the omitted group.

There are two extensions: Firstly, the fixed effects model can be extended to include a time-specific effect as well. One way to formulate this is simply to add the time effect by the inclusion of an additional (T-1) dummies. This is implemented in our empirical work in the following section. Secondly, if we have missing data in the survey we have what we call an unbalanced panel. With an unbalanced panel the full sample size is $\sum_{i=1}^{n} T_i$ instead of nT, where T_i is the number of observations for the ith group. Thus group means must be based on T_i, which varies across groups.

A major limitation of within estimation is that the coefficients of time-invariant regressors are not identified in the within model. Thus in our panel earnings regressions we are not able to estimate the gender or regional impact. Random effects estimators do however permit estimation of coefficients of time-invariant regressors. This is an important point, fixed effects only allows the case of time-variant explanatory variables.

4.4.2. Random Effects

If the individual effects are uncorrelated with the regressors, then it might be appropriate to model the individual-specific constant terms as randomly distributed across cross-sectional units. This view would be appropriate if we believed that sampled cross-sectional units were drawn from a large population. The gain to this approach is that it substantially reduces the number of parameters to be estimated. The cost is the possibility of inconsistent estimates should the assumption be inappropriate. Rearranging equation (4.10), we obtain:

$$Ln(Y_{it}) = \alpha + \beta X_{it} + u_i + \varepsilon_{it} \tag{4.13}$$

There is now a single constant term which is the mean of the unobserved heterogeneity, α=E($Z_i\pi$), and is constant over time. u_i (u_i={$Z_i\pi$–E[$Z_i\pi$]}) is the random heterogeneity specific to the i'th observation. We assume further that:

$$\begin{aligned}
&E[\varepsilon_{it} | X] = E[u_i | X] = 0,\\
&E[\varepsilon_{it}^2 | X] = \sigma_\varepsilon^2,\\
&E[u_i^2 | X] = \sigma_u^2,\\
&E[\varepsilon_{it}\, u_j | X] = 0 \quad \text{for all } i, t \text{ and } j,\\
&E[\varepsilon_{it}\, \varepsilon_{js} | X] = 0 \quad \text{if } t \neq s \text{ or } i \neq j,\\
&E[u_i\, u_j | X] = 0 \quad \text{if } i \neq j.
\end{aligned} \tag{4.14}$$

Based on these assumptions, random effects model is performed by Generalised Least Squares (GLS)[63]. GLS based on the true variance component is a best linear unbiased estimator (BLUE) and all the feasible GLS estimators considered are asymptotically efficient when the number of observations is large.

The crucial distinction between fixed effects estimation and random effects estimation is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model. There are two key tests to discuss. Firstly, Breusch and Pagan (1980) have devised a Lagrange multiplier (LM) test for the random effects model based

[63] Details of GLS see Greene (2003) pp.293-98.

on the OLS residuals. The null hypothesis of the test is that the variance of the individual effects equals zero against the alternative that it does not equal zero. Under the null hypothesis, LM is distributed as chi-squared with one degree of freedom. If we obtain a LM test statistic which exceeds the (for example, 95 percent) critical value, we conclude that the classical regression model with a single constant term is inappropriate, thus the result of the test is to reject the null hypothesis in favour of the random effects model. Secondly, the specification test devised by Hausman (1978) is used to test for whether the random effects are independent of the regressors. This is a general test to compare any two estimators. The test is based on the idea that under the hypothesis of no correlation, both fixed effects model and GLS are consistent, but the fixed effects model is inefficient, whereas under the alternative, the fixed effects model is consistent, but the GLS one is not. Therefore, under the null hypothesis, the two estimates should not differ systematically. To sum up, if based on a LM test, which is decisive that there are individual effects, and a Hausman test, which suggests that these effects are uncorrelated with the regressors in the model, we could conclude that of the two alternatives we have considered, the random effects model is the better choice. However, it is not possible to use the Hausman test to compare fixed and random effects for a model with dummy variables such as gender and region which are time-invariant.

4.5. Empirical Results

4.5.1. Changes in Returns to Education over Time

We start by estimating equation (4.2) for each year respectively, and the coefficient estimates are reported in Table 4-6[64]. The results show that education is significant with the expected signs in all four years. Earnings increase with education, and this positive linkage is consistent with the research that has previously been done on this subject in China. In 1989 the marginal return is 4%[65]. This percentage declined from 1989 to 1993, and increased thereafter. Note that it has noticeably increased from 4% in 1997 to 7% in 2000. Still, this is somewhat lower than some studies in China and studies in other developing countries as mentioned in the literature, and it may appear to be too low to be consistent with China's recent development. However, this is consistent with the long-term egalitarian income distribution principles pursued by the Chinese government in urban areas, and further it may partly reflect the underdevelopment in China's market economy in the transition period since economic reform is a continuous process and most policies take time to produce any changes. One reason for our relatively low estimates is that the dependant variable includes subsidies and bonuses which tend to flatten the distribution. To a large extent, our estimates reflect the income returns to education, but not the earnings, net of benefits, returns which may be higher.

In addition, we also need to be concerned by the fact that we could not separate the rural-to-urban migrants from urban residents in our sample when estimating returns to education in urban China. As with many other studies the data was simply not available to allow us to do this. As mentioned in chapter 2, beginning with the reform period in the late 1970s and accelerating during the late 1990s, reforms to the household registration system have been implemented by Chinese authorities. Restrictions on rural-urban labour

[64] The reported coefficients were estimated by STATA 9.0.

[65] The estimated change in earnings Y_i for individual i is calculated by (e^{β_i} –1).

migration, as well as on obtaining urban residence permits for rural migrants, have been gradually relaxed. According to Brooks and Tao (2003), the National Bureau of Statistics estimates there were about 80 million permanent rural-to-urban migrants (i.e., those living in urban areas for more than six months) between 1990 and 2000. Therefore, our estimates include a sample of rural-to-urban migrants that are likely to have a higher return to education[66] (Meng and Zhang, 2001). Whether this is still true for current migrants is an empirical question that is difficult to answer since there are still too few studies on these issues given the available data.

In order to test the robustness of the estimated increasing trend in the returns to education and to investigate whether the inclusion of different sectors[67] changes the estimated coefficients of education, we add two sectoral dummies: state sector and private sector. People not in either of these sectors include, e.g., those in collectives. Table 4-7 shows that adding sectoral dummies reduced the estimated returns to education for 1989, however, it has a negligible effect in 1997 and 2000. In addition, the trend of increasing returns to education over time is robust to the inclusion of these sectoral dummies, except that the coefficient in 1993 fails to be significant. Hence, we conclude that in general rising returns to education do not specifically reflect sectoral changes in the labour

[66] Meng and Zhang (2001) found that the rate of return to education is around 1% higher for rural migrants than for urban residents.

[67] In addition to the inclusion of different sectors, we have also added occupation variables in the regressions. Adding occupation dummies has a negligible effect on the coefficients of education in 1997 and 2000, but in 1989 and 1993, the coefficients of education fail to be significant. The latter to some extent is consistent with the discussions in the literature that a high degree of multicollinearity exists between education and occupation (Li, 2003). Psacharopoulos and Patrinos (2002) also argued that the inclusion of occupation in the Mincer earnings function leads to ‘stealing’ part of the effect of education on earnings that comes from occupational mobility. Moreover, it seems also likely that occupations and wages are jointly determined, in other words, simultaneity may exist between them. We do not report results here.

market.

Table 4-6 Wages Returns to Education in Urban China – Basic Equations

Explanatory Var.	**Ln(Wages)**			
	1989	1993	1997	2000
Education	0.042***	0.020**	0.042***	0.064***
	(4.87)	(2.57)	(6.46)	(8.89)
Experience	0.098***	0.076***	0.032***	0.010
	(15.09)	(11.97)	(5.61)	(1.49)
Expsq	-0.002***	-0.002***	-0.001***	-0.00003
	(-15.97)	(-14.25)	(-4.87)	(-0.22)
Gender	0.289***	0.266***	0.222***	0.143***
	(5.54)	(5.65)	(6.07)	(3.54)
Liaoning	-0.092	-0.530***	-	-0.302***
	(-0.94)	(-6.16)		(-3.66)
Heilongj	-	-	-0.102	0.022
			(-1.48)	(0.29)
Shandong	0.026	-0.641***	-0.279***	-0.092
	(0.26)	(-6.55)	(-4.11)	(-1.12)
Henan	-0.132	-0.571***	-0.312***	-0.303***
	(-1.32)	(-5.93)	(-4.85)	(-4.09)
Hubei	-0.116	-0.340***	-0.209***	-0.344***
	(-1.11)	(-3.74)	(-3.47)	(-4.81)
Hunan	-0.043	-0.287***	-0.140**	-0.094
	(-0.42)	(-3.17)	(-2.26)	(-1.34)
Guangxi	-0.528***	-0.541***	-0.310***	-0.444***
	(-5.14)	(-5.63)	(-4.25)	(-5.30)
Guizhou	-0.899***	-1.199***	-0.504***	-0.293***
	(-8.18)	(-12.62)	(-6.76)	(-3.53)
R-squared	0.238	0.268	0.136	0.15
Adj R-squared	0.233	0.263	0.128	0.14
No. of the Obs.	1852	1496	1178	1063

Notes: t statistics in parentheses, *** Significant at 1%; ** Significant at 5%; * Significant at 10%.

Table 4-7 Wages Returns to Education in Urban China – adding Sectors

Explanatory Var.	Ln(Wages)			
	1989	1993	1997	2000
Education	0.023***	0.012	0.036***	0.056***
	(2.60)	(1.51)	(5.40)	(7.28)
Experience	0.092***	0.075***	0.034***	0.009
	(14.25)	(11.82)	(6.06)	(1.30)
Expsq	-0.002***	-0.002***	-0.001***	-0.00003
	(-15.86)	(-14.39)	(-5.22)	(-0.20)
Gender	0.289***	0.265***	0.231***	0.149***
	(5.61)	(5.65)	(6.46)	(3.71)
State	0.384***	0.157***	0.213***	0.176***
	(6.37)	(2.76)	(5.02)	(3.54)
Private	-0.277***	-0.166	0.448***	0.094
	(-2.68)	(-1.61)	(7.15)	(1.55)
Liaoning	-0.207**	-0.588***	-	-0.351***
	(-2.12)	(-6.68)		(-4.21)
Heilongj	-	-	-0.167**	-0.029
			(-2.43)	(-0.38)
Shandong	-0.044	-0.667***	-0.283***	-0.102
	(-0.44)	(-6.82)	(-4.27)	(-1.23)
Henan	-0.190*	-0.619***	-0.335***	-0.317***
	(-1.86)	(-6.37)	(-5.28)	(-4.30)
Hubei	-0.206**	-0.401***	-0.241***	-0.375***
	(-1.98)	(-4.36)	(-4.03)	(-5.19)
Hunan	-0.074	-0.308***	-0.197***	-0.113
	(-0.72)	(-3.40)	(-3.24)	(-1.61)
Guangxi	-0.537***	-0.549***	-0.393***	-0.437***
	(-5.15)	(-5.70)	(-5.45)	(-5.25)
Guizhou	-0.966***	-1.232***	-0.567***	-0.314***
	(-8.63)	(-12.50)	(-7.69)	(-3.80)
R-squared	0.267	0.276	0.176	0.160
Adj R-squared	0.262	0.269	0.167	0.148
No. of the Obs.	1852	1496	1178	1063

Notes: t statistics in parentheses, *** Significant at 1%; ** Significant at 5%; * Significant at 10%.

Table 4-8 Wages Returns to Education in Urban China – adding Interaction Terms

Explanatory Var.	Ln(Wages)			
	1989	1993	1997	2000
Education	0.031^{**}	0.013	0.036^{***}	0.043^{***}
	(2.16)	(0.82)	(2.83)	(2.74)
Experience	0.092^{***}	0.077^{***}	0.036^{***}	0.008
	(14.15)	(12.13)	(6.36)	(1.15)
Expsq	-0.002^{***}	-0.002^{***}	-0.001^{***}	-0.00002
	(-15.79)	(-14.70)	(-5.53)	(-0.13)
Gender	0.711^{***}	0.593^{***}	0.432^{***}	0.239^{*}
	(6.28)	(4.81)	(3.87)	(1.77)
Genedu	-0.053^{***}	-0.036^{***}	-0.021^{*}	-0.009
	(-4.17)	(-2.83)	(-1.89)	(-0.72)
State	0.185	0.025	0.118	-0.104
	(1.50)	(0.18)	(0.94)	(-0.64)
Staedu	0.026^{*}	0.018	0.012	0.028^{*}
	(1.80)	(1.13)	(0.90)	(1.81)
Private	-0.725^{***}	-1.144^{***}	0.039	0.001
	(-3.58)	(-4.29)	(0.21)	(0.01)
Priedu	0.064^{**}	0.118^{***}	0.046^{**}	0.010
	(2.47)	(3.95)	(2.42)	(0.52)
Liaoning	-0.209^{**}	-0.592^{***}	-	-0.354^{***}
	(-2.15)	(-6.75)		(-4.24)
Heilongj	-	-	-0.166^{**}	-0.042
			(-2.41)	(-0.55)
Shandong	-0.063	-0.672^{***}	-0.283^{***}	-0.093
	(-0.63)	(-6.89)	(-4.27)	(-1.12)
Henan	-0.204^{**}	-0.622^{***}	-0.340^{***}	-0.324^{***}
	(-2.01)	(-6.42)	(-5.37)	(-4.38)
Hubei	-0.217^{**}	-0.406^{***}	-0.241^{***}	-0.367^{***}
	(-2.10)	(-4.42)	(-4.05)	(-5.08)
Hunan	-0.077	-0.315^{***}	-0.199^{***}	-0.106
	(-0.75)	(-3.47)	(-3.28)	(-1.50)
Guangxi	-0.540^{***}	-0.571^{***}	-0.397^{***}	-0.431^{***}
	(-5.18)	(-5.94)	(-5.51)	(-5.16)
Guizhou	-0.974^{***}	-1.221^{***}	-0.580^{***}	-0.318^{***}
	(-8.74)	(-12.37)	(-7.88)	(-3.84)
R-squared	0.275	0.287	0.184	0.163
Adj R-squared	0.269	0.280	0.172	0.149
No. of the Obs.	1852	1496	1178	1063

Notes: t statistics in parentheses, *** Significant at 1%; ** Significant at 5%; * Significant at 10%.

Further, we examine whether the changes in the returns to education differ systematically by gender and sector. The coefficients of the interactions between education and gender, education and sector are shown in Table 4-8. The results indicate that men earn more than women in all four years and hence there is a large gender gap. The size of the coefficients suggests that the gap between men's wages and women's declined steadily from 1989 to 2000. However, education modifies this basic gender gap. From the coefficient on the interaction term gender*education, we can see that the returns to education are lower for men than women, which is consistent with the literature. The size of the coefficient declines from 1989 to 2000, which suggests that the impact of education on the gender gap is also declining. Given that the gender gap itself is decreasing this is perhaps not too surprising. There are several possible explanations in the literature as to why educational returns are higher for women. Deolalikar (1993) argued that men have a comparative advantage in physical strength so that education becomes relatively more important to women who focus on more skill-intensive jobs. In addition, the sample distribution by genders in Table 4-5 shows that the average educational obtainments for women are less than for men, although the gap is less evident in 2000. This is to some extent consistent with the discussion that fewer women achieve high levels of education, which reduces the relative supply of high skilled women (Li, 2003). The gender gap may be less evident as we move up the salary, and hence occupational, scale. In other words, the gender gap is possibly partially caused by women, particularly those with relatively low levels of education, being crowded into specific and low paid occupations. As argued by Chow (1995), higher educational qualifications result in women being able to compete in occupations with men, escaping these low paid occupations.

To test the robustness of the results, we start by running separate regressions for men and women. As is shown in the odd columns of Table 4-9, returns to education are higher for women in 1989, 1993 and 1997, and in 2000, it is similar for men and women. These results, as expected, are consistent with the results shown in Table 4-8 with the inclusion of an interaction term between gender and education. However, as discussed in the labour supply literature, such estimates may be biased if the sample of people for whom we observe earnings is not randomly selected. More specifically, it is possible that the higher returns to education for women is due to greater self-selection of women into the labour force relative to men, whose labour force participation is nearly universal (Zhang, *et al*., 2005). On average, the proportion of men participating in work is approximately 9% higher than for women in all four years[68], although this difference declined from 1989 to 2000. To test this hypothesis directly, we estimate a Heckman selection model for earnings. This included marital status and number of household members[69] in identifying the selection equation. The results are shown in the even columns of Table 4-9. For women, the Mills' ratios from the selection equations are not significant in the wage equations in all four years, and similar results applied for men, apart from 1989. Thus this source of bias may not be too important in the (urban) Chinese context, and therefore, this effect is not sufficient to explain the higher returns to education for women. In part this

[68] These numbers are consistent with the figures published by Labour and Social Trends in Asia and the Pacific 2005. For men, the labour force participation rate (aged 15-64) is 89.6%, 90.0% and 89.5% in 1990, 1995 and 2000 respectively. For women, the proportion is 79.9%, 80.3% and 80.0% respectively.

[69] The assumption regarding the number of household members is that the more of the members in household, the less participation by women due to the increased requirements for either childcare or eldercare. It has been largely used in identifying the selection equation in Chinese context (see, e.g., Li, *et al*., 2005; Zhang, *et. al*., 2005).

may be because the gap noted above between male and female participation rates is not that great compared to many developed countries, particularly in the recent past.

Given the mixed economies in contemporary China, it seems logical to compare differences across sectors based on their level of closeness to the market. Table 4-7 shows that individuals working in the private sector earn less than those in the state and collective sectors, although this trend changes in 1997. When we further include state*education (staedu) and private*education (priedu) in the regression, Table 4-8 shows the returns to education have been consistently higher for individuals in the private sector from 1989 to 1997 when the principles of human resource allocation are more governed by market mechanisms. Note also that returns to education in the private sector increased from 1989 to 1993, and declined significantly thereafter. In 2000, the higher rate of returns to education has shifted to individuals in the state sector. This may in part be due to substantial layoffs, retirements, and exits from SOEs since 1997, since when the connection between education and the corresponding reward system has been more visible in the state sector. One possible concern with this result is that the range of the private sector in the survey is defined too broadly as shown in Table 4-2. For example, domestic private enterprises and foreign invested enterprises both belong to the category of the private sector in the survey, but their characteristics may be sufficiently different to warrant separate treatment (Zhao, 2002). This may in part lead to the changes in the returns to education in the private sector.

Table 4-9 Wages Returns to Education by Gender

Explanatory Var.	Ln(Wages)_Female							
	1989		1993		1997		2000	
	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman
Education	0.030**	0.028	0.018**	-0.003	0.046***	0.039**	0.051***	0.055**
	(2.31)	(0.86)	(2.01)	(-0.10)	(4.27)	(2.21)	(4.01)	(2.26)
Experience	0.091***	0.092***	0.058***	0.067***	0.034***	0.036***	-0.0004	0.0003
	(10.57)	(7.75)	(8.91)	(6.55)	(3.79)	(3.57)	(-0.04)	(0.96)
Expsq	-0.002***	-0.002***	-0.001***	-0.002***	-0.001***	-0.001***	0.0002	0.0001
	(-12.74)	(-11.26)	(-11.28)	(-9.31)	(-3.28)	(-3.18)	(0.79)	(0.48)
State	0.356***	0.424***	0.021	0.230***	0.256***	0.253***	0.215***	0.264***
	(4.45)	(4.74)	(0.36)	(2.64)	(3.88)	(3.80)	(2.71)	(3.32)
Private	-0.665***	-0.493***	0.515***	-0.456***	0.286***	0.293***	0.092	0.126
	(-4.17)	(-2.80)	(4.44)	(-2.80)	(3.06)	(3.11)	(0.96)	(1.26)
Regional Dummies	yes	yes	yes	Yes	yes	yes	yes	yes
Married		-0.324**		-0.310*		-0.440**		-0.420**
		(-2.43)		(-1.89)		(-2.54)		(-2.30)
No. of Household Members		-0.015		-0.134***		-0.127***		-0.023*
		(-0.65)		(-3.86)		(-3.23)		(-1.65)
No. of Household Members^2		-0.0002		-0.001***		-0.001***		-0.0002
		(-0.63)		(-3.64)		(-3.11)		(-1.28)
λ (Mills ratio)		0.178		-0.271		-0.079		-0.013
		(0.60)		(-1.05)		(-0.49)		(-0.06)
R-squared	0.399		0.328		0.161		0.139	
Adj R-squared	0.390		0.315		0.141		0.114	
Wald $\chi^2(22)$		623.02		501.22		301.74		276.67
No. of the Observations	810	1351	648	980	514	937	452	771

Table 4-9 continued

Explanatory Var.	**Ln(Wages)_Male**							
	1989		1993		1997		2000	
	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman
Education	0.024**	-0.016	0.007	0.006	0.029***	0.039***	0.052***	0.055**
	(2.15)	(-0.88)	(0.67)	(0.28)	(3.39)	(2.73)	(4.94)	(2.22)
Experience	0.096***	0.077***	0.078***	0.071***	0.032***	0.036***	0.011	0.012
	(11.11)	(6.66)	(9.02)	(7.31)	(4.36)	(4.42)	(1.16)	(1.00)
Expsq	-0.002***	-0.002***	-0.002***	-0.002***	-0.001***	-0.001***	-0.0001	-0.0001
	(-11.20)	(-8.01)	(-10.27)	(-8.83)	(-3.68)	(-3.80)	(-0.34)	(-0.36)
State	0.352***	0.326***	0.118	0.150*	0.193***	0.184***	0.119*	0.119*
	(4.26)	(3.94)	(1.51)	(1.75)	(3.47)	(3.24)	(1.69)	(1.72)
Private	-0.203	-0.200	0.017	-0.090	0.590***	0.620***	0.090	0.089
	(-1.57)	(-1.58)	(0.12)	(-0.62)	(6.93)	(7.04)	(1.06)	(1.06)
Regional Dummies	yes	yes	yes	Yes	yes	yes	yes	yes
Married		-0.352**		-0.214		-0.217		0.078
		(-2.33)		(-1.18)		(-1.22)		(0.69)
No. of Household Members		-0.246**		-0.066*		-0.873***		-0.013
		(-2.38)		(-1.93)		(-4.93)		(-1.13)
No. of Household Members^2		0.015*		-0.001*		0.084***		-0.0001
		(1.67)		(-1.64)		(4.53)		(-1.11)
λ(Mills ratio)		-0.722***		0.058		0.144		0.038
		(-3.04)		(0.18)		(0.89)		(0.12)
R-squared	0.205		0.227		0.165		0.123	
Adj R-squared	0.195		0.216		0.150		0.104	
Wald $\chi^2(22)$		359.29		359.14		336.99		235.62
No. of the Observations	1019	1372	818	958	664	1031	611	827

Notes: t statistics in parentheses, *** Significant at 1%; ** Significant at 5%; * Significant at 10%.

It is clear that returns to education differ significantly between genders and sectors as shown in Table 4-8, thus it is necessary to note that estimated returns to education are not directly comparable due to the selection of different samples and the inclusion of different control variables in model specifications. Apart from this, we will also look at the earning returns to experience and different regions, which also show strongly significant effects in the earnings equation. From the coefficients of experience and experience squared[70] in Table 4-8, we can see that the effect of potential experience on earnings is convex, first increasing and then decreasing after reaching a given number of years. The highest return for years of experience varies between 24 (in 1993) and 29 (in 1997), and it fails to be significant in 2000. Thus for someone leaving education at age 18, these peaks would be at age 42 and 47 respectively. These curves are illustrated in Figure 4-2 from which it is noticeable that the curves have been systematically getting flatter. For example, wages' returns for someone with fifteen-year experience decline significantly from 1989 to 1997. This may partly reflect a declining impact of seniority and an increasing role of market forces, which are both closely related to China's economic reforms.

To further examine whether this declining impact of experience on earnings over time is restricted across different ages, we split our sample into two age cohorts: the young – ages 18-35, and the middle aged and above – ages 36-65[71]. From Table 4-10, we can see

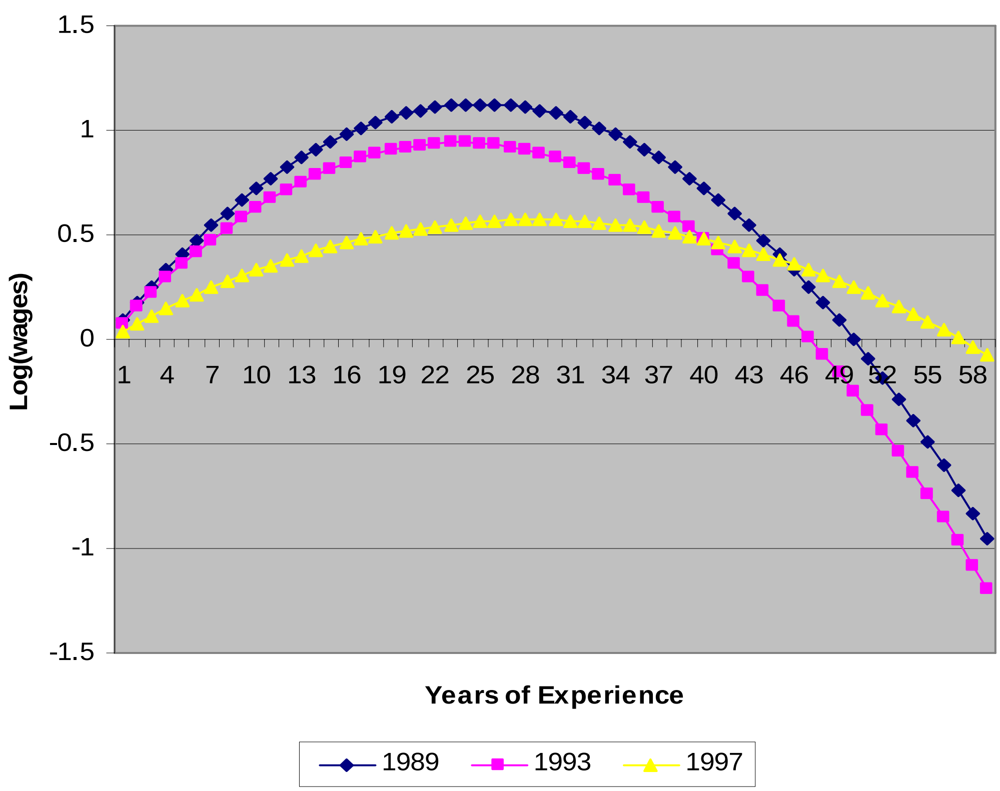
[70] The coefficients of experience and experience-squared are extremely close in Table 4-6, Table 4-7 and Table 4-8.

[71] Note that we have also tried to separate the cohorts of the 'middle aged and above' into the 'middle aged (36-50)' and the 'older (51-65)', however, we do not find any significant differences between them for either experience or education respectively. We do not report results here.

that experience has a higher first-order impact for the young than the middle aged and above in all four years. This is consistent with the discussions by Mehta and Villarreal (2004). Experience provides on the job skill acquisition, thus it may correlate with obsolescence. In addition, the greatest rate of skill acquisition will come in the early years of employment. Thus the young experience high rates of skill acquisition and limited obsolescence effects, while comparatively older workers may experience lower marginal rates of skill acquisition and greater obsolescence, reducing the continuing impact of experience on their earnings. Indeed, as already emphasised, we find this for all four years. In addition, Table 4-10 also shows that the rate of returns to education diminishes for the middle aged and above compared to the young. Similar to the impact of experience, possible explanations for this difference include an education vintage effect which may reflect a higher return for more recently obtained education. We will discuss this education vintage later. Note that it is worth emphasising that we have relatively few observations in these regressions in Table 4-10.

Regional impacts are estimated using separate dummy variables in the earnings equation. We take Jiangsu, one of the developed coastal provinces, as the reference province where the economic transition is comparatively advanced compared to the other provinces. The results in Table 4-8 show that almost all regional coefficients are significantly negative in all four years, which indicates that Jiangsu has the highest earning returns in China. This is also consistent with the focus of China's economic reforms on mainly developing the coastal regions in the southern and eastern parts based on their readily accessible geographic location (Huang, *et al.*, 2003). The results show that in 1989, individuals' earnings in Guizhou are about 37.9% of those in Jiangsu with assumed similar characteristics (education, gender, etc.). Although people in different provinces may not have on average similar characteristics, this still leads us to believe that people in Guizhou will be much more poorly paid than those in Jiangsu. However, there is a declining trend of the earnings differences, specifically between Jiangsu and two less-developed western provinces, Guangxi and Guizhou. This indicates that the earnings gap between those provinces is getting smaller, and people' earnings in Guizhou increase to 72.6% of those in Jiangsu in 2000, other things being equal [72].

Figure 4-2 The Impact of Experience



Source: Table 4-8
Notes: We do not include the year 2000 in this figure as it fails to be significant in the estimate.

[72] To the extent that shadow economy activity is a problem in China, these figures may also reflect differences between regions in shadow economy activity.

Table 4-10 Wages Returns to Experience by Age Cohorts

Explanatory Var.	**Ln(Wages)**							
	1989		1993		1997		2000	
	15-35	36-65	15-35	36-65	15-35	36-65	15-35	36-65
Education	-0.004	0.029**	0.046***	-0.004	0.047***	0.027***	0.057***	0.047***
	(-0.26)	(2.45)	(2.86)	(-0.44)	(3.72)	(2.95)	(3.57)	(4.56)
Experience	0.154***	0.094***	0.088***	0.056***	0.027	0.018	0.051**	0.022
	(7.54)	(3.96)	(3.51)	(2.81)	(1.37)	(1.00)	(2.07)	(0.84)
Expsq	-0.005***	-0.002***	-0.002*	-0.001***	-0.0002	-0.0004	-0.002*	-0.0002
	(-5.58)	(-5.69)	(-1.79)	(-5.32)	(-0.23)	(-1.44)	(-1.73)	(-0.39)
Gender	0.100	0.485***	0.115*	0.397***	0.245***	0.232***	0.128*	0.144***
	(1.42)	(6.51)	(1.65)	(6.32)	(4.81)	(4.70)	(1.92)	(2.63)
State	0.510***	0.190**	0.046	0.218***	0.147**	0.269***	0.113	0.182***
	(6.34)	(2.13)	(0.56)	(2.77)	(2.47)	(4.53)	(1.36)	(2.71)
Private	0.024	-0.750***	-0.054	-0.359**	0.469***	0.377***	0.112	0.049
	(0.18)	(-4.74)	(-0.39)	(-2.36)	(5.80)	(3.94)	(1.23)	(0.54)
Regional Dummies	Yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.181	0.361	0.164	0.356	0.236	0.147	0.118	0.166
Adj R-squared	0.169	0.352	0.146	0.346	0.214	0.130	0.089	0.147
No. of the Obs.	912	940	624	872	475	703	451	612

Notes: t statistics in parentheses, *** Significant at 1%; ** Significant at 5%; * Significant at 10%.

4.5.2. Panel Data Analysis

We further estimate the pooled data set where the panel is an unbalanced one. We begin by estimating equation (4.2) with the inclusion of sectoral dummies using OLS, FE and RE estimators. The individual effects are statistically significant through the *F*-test/ BPLM Test, thus justifying the use of panel methods. As mentioned in section 4.4, the FE and RE models allow for specific individual effects. The RE estimator treats the individual effects as being randomly distributed across individuals though, unlike the fixed effects procedures, it requires the individual effects to be independent of the error term. This assumption is tested using Hausman test. However, it is clear that the test is performed only with the coefficients of the time-varying variables included in both models. The inclusion of regional dummy variables as well as gender precludes the use of fixed effects per se (Bortolotti, *et al.*, 2004). Regional dummies are crucial in the earnings equation as they capture the group differences by different regions. Statistically, the estimates of regional dummies are almost all significant at 1% level and Wald-tests further indicate that those as a group significantly explain much of the variation in earnings. Based on these, we will present only the results of the random-effects, and the overall significance of the model is established by the overall R-squared. The results are shown in Table 4-11.

Table 4-11 Wages Returns to Education in Urban China – Results of Random-effects

Explanatory Var.	Ln(Wages)			
	RE-reg1	RE-reg2	RE-reg3	RE-reg4
Education	0.022***	0.027***	0.046***	0.012
	(4.03)	(4.94)	(4.15)	(0.76)
Experience	0.068***	0.069***	0.081***	0.081***
	(17.20)	(17.30)	(11.31)	(11.60)
Expsq	-0.001***	-0.001***	-0.002***	-0.002***
	(-20.14)	(-20.07)	(-16.69)	(-16.93)
Expedu			-0.0007**	-0.0008**
			(-1.97)	(-2.27)
Gender	0.271***	0.259***	0.262***	0.498***
	(8.07)	(7.60)	(7.69)	(6.46)
Genedu				-0.027***
				(-3.46)
State	0.250***	0.242***	0.239***	0.191***
	(8.28)	(7.99)	(7.89)	(2.61)
Staedu				0.006
				(0.79)
Private	0.071	0.064	0.062	-0.17
	(1.63)	(1.46)	(1.41)	(-1.52)
Priedu				0.037***
				(3.08)
Liaoning	-0.358***			
	(-5.44)			
Heilongj	-0.130			
	(-1.47)			
Shandong	-0.243***			
	(-3.71)			
Henan	-0.355***			
	(-5.55)			
Hubei	-0.300***			
	(-4.66)			
Hunan	-0.196***			
	(-3.06)			
Guangxi	-0.607***			
	(-9.16)			
Guizhou	-0.929***			
	(-13.60)			

Table 4-11 continued

Explanatory Var.	Ln(Wages) RE-reg1	RE-reg2	RE-reg3	RE-reg4
Distbeij		-0.0002**	-0.0002**	-0.0004**
		(-2.30)	(-2.26)	(-2.33)
Distshan		-0.0006***	-0.0006***	-0.0009***
		(-8.01)	(-7.95)	(-4.92)
Dbeiedu				0.00003**
				(1.97)
Dshanedu				0.00004**
				(2.23)
Year 1989	-1.006***	-1.025***	-1.023***	-1.007***
	(-29.65)	(-30.53)	(-30.51)	(-30.81)
Year 1993	-0.634***	-0.657***	-0.655***	-0.640***
	(-19.70)	(-20.61)	(-20.55)	(-20.49)
Year 1997	-0.330***	-0.330***	-0.330***	-0.326***
	(-11.10)	(-11.06)	(-11.05)	(-11.08)
Overall R-sq	0.356	0.347	0.346	0.352
No. of the Obs.	5589	5589	5589	5589

Notes: t statistics in parentheses, *** Significant at 1%; ** Significant at 5%; * Significant at 10%. Regressions estimated based on panel data with random-effects (RE-reg). RE-reg1 is the basic equation with regional dummies. RE-reg2 replaces the regional dummies to instances to Beijing and Shanghai. RE-reg3 further includes the interaction term experience*education. RE-reg4 is the full equation with all interactive variables included.

The first two columns of Table 4-11 present results with both the inclusion of the regional dummy variables and the distance of the province from both Beijing (Distbeij) and Shanghai (Distshan) – although this made no noticeable difference to the other coefficients in this regression. Beijing is the capital of China and the country's political, economic, cultural and transportation centre. Shanghai is the biggest port in China and is the most important centre of trade and finance and also information and culture. These two largest cities account for a large part of the variation in regional incomes as argued in chapter 2. We expect that the further a region is from either of these two centres, the less developed it will be with possible implications for earnings. In addition, Beijing, being the political centre, may also have an impact which to some extent inhibits economic development, or more precisely, its policies specifically on opening to the outside world are not as wide as Shanghai's. Based on this, we would expect the more significant linkage to be between the distance to Shanghai and earning implications. The overall R-squared indicates that these two variables largely explain the regional differences identified by the regional dummies. Note that the distances are calculated based on the capitals of the nine provinces to these two metropolises, so they might have some limitations which restrict the full explanation of the regional differences[73]. Note too that this is consistent with recent work in economic geography which emphasizes the importance of proximity to large population mass. However, our results go further than this and emphasize the importance of not simply mass but proximity to the two large power centres in China, one economic, and the other political.

The results indicate that distance from Shanghai is significant at 1% level and its negative coefficient indicates that there is an inverse relationship between individuals' earnings and the distance from where they work to Shanghai. Less pronounced, but still significant at the 5% level, is the negative coefficient of distance to Beijing. This means that during the transition period Shanghai has become the most 'profitable' economic centre in China. People earn less, the further they are far from Beijing, and more noticeable, Shanghai. This is consistent with our expectations. Economic reform and the opening up to the outside world made coastal regions the main development focus, and they have been much more able to attract FDI based on their locations. For example, Shanghai's FDI

[73] We have done the Cox-Pesaran test for these two (non-nested) models based on pooled data set, and results indicate that we could not reject the model specification with regional dummies.

volume in 2000 totalled 3.1 billion US dollars, 87.7% higher than Beijing's[74]. Moreover, international migration in Shanghai now runs at the highest level in China.

The third column of Table 4-11 aims to examine whether there is a possible vintage effect of education by the inclusion of the interaction term education*experience (eduexp) as argued by section 3. The essence of this exercise is to justify whether the impact of education on earnings declines as the time since the individual left formal education increases. We have hypothesized that education improves people's skills, marginal product and hence their wages, but, that in an evolving labour market, the value of education declines with years since the individual left education. This may be so because skills date. Certainly at the beginning of the second millennium this is truer than even say 50 years ago. The coefficient on education*experience is significantly negative, at the 5% level. This indicates that returns to education decline with time, i.e., the greater the period since people left formal education, the lower are the returns to education. This is consistent with the hypothesis that in part at least education's impact on earnings is by skill enhancement and it is not simply a proxy for ability. Assuming a working life of forty years, the coefficients suggest that approximately 61% of the returns to education are eroded during a working life and this provides a lower bound to the actual impact of education itself. Another way to take into account this possible vintage effect is to run separate regressions for different age cohorts (or years of work experience) as we have implemented in the last section. We split our sample into two age cohorts: the young – ages 18-35, and the middle aged and above – ages 36-65. These two methods give very similar results. The rate of return to education diminishes for the middle aged and above

[74] Data come from China Statistical Yearbook, 2001.

(more experienced), and reflect a higher return for more recently obtained education. We do not report results in Table 4-11.

The final column of Table 4-11 shows the results with the inclusion of the interactions between education and gender, experience, sector and distance to Beijing (dbeiedu) and Shanghai (dshanedu). The results show generally similar results on returns to education restricted by genders and sectors. Specifically, people earn less when they work further away from Shanghai and Beijing, however, they do have higher returns to education. This is to some extent consistent with Li (2003) and Zhang, *et al.*'s (2005) results which people have higher returns to education in less developed regions. Therefore, there is no single rate of return to education, and in a sense discussion on such 'a return' is misleading. For example, the rate of return to education (R_e) for a man working in the private sector with 15 years of working experience in Jiangsu province is 3.4%, calculated as follows:

$$
\begin{aligned}
R_e = {} & \exp [0.012 - 0.027*1(\text{man}) - 0.0008*15(\text{experience}) + 0.037*1(\text{private sector}) \\
& + 0.00003*560(\text{distance to Beijing}) + 0.00004*166(\text{distance to Shanghai})] - 1 \\
= {} & 0.034 \qquad (4.15)
\end{aligned}
$$

Finally, the panel regressions also include year fixed effects which are all significant at the 1% level. The coefficients, which are reported in Table 4-11, suggest a steady increase from 1989 to 2000.

4.6. Conclusions

In this chapter, we have used the panel data set including 1989, 1993, 1997 and 2000 from the China Health and Nutrition Survey, to estimate the rate of private returns to education in urban China. The explanatory variables include education, experience and experience squared, as in the Mincer earnings function, but also gender, sector, region, distance to Beijing and Shanghai, and education-interaction terms. This is the first time this data source has been used in this context, although to a considerable extent our results are consistent with those of others for other countries and also with the limited amount of work that has been done on China.

The estimated rate of return to education based on our baseline model is 4%, 2%, 4% and 7% in 1989, 1993, 1997 and 2000 respectively. These results show that there is a noticeable increase, specifically from 1997 to 2000. In addition, the return depends on gender and also indeed sectors and regions. This is an important point, and there is no single rate of return to education. As with many other countries education reduces the gender earnings gap. Females earn less than males in all four years, but the gap appears to have been steadily declining. Again as with many other countries, education further reduces this basic gap, but just as the gap has been declining so the marginally greater return to education for women has also been declining. Nonetheless the existence of this difference in the returns to education both in China and indeed elsewhere suggests a greater incentive for women to pursue educational study, and to an extent one wonders why they do not, although that is the subject of a different study to ours.

One of the basic problems facing this type of research is the problem of whether education enhances ability or merely reflects ability. If the latter, the estimated impact of education on earnings is no more than correlation. However, we have argued that if the former is true then in a changing world, skills, including educationally enhanced skills, will date and enhance the value of education in enhancing earnings should decline as the period since formal education completed increases. We found strong evidence for this in the regression results and hence conclude that education enhances ability and does not simply reflect it. Once more we emphasize that this view depends on the assumption that the value of education as a signal does not decline with age. Still, controversies exist concerning whether the omitted ability variables bias the OLS estimates, as well as concerning the extent of the potential bias. Our analysis has added to this literature, but it does not by itself fully resolve the issue. It is an extra piece to the jigsaw and with the other pieces helps us see the complete picture.

Another key result is the extent of regional differences and the impact of distance from Beijing and Shanghai on earnings. People earn less, the further they are far from Beijing, and more noticeably, Shanghai; however, they do enjoy higher rate of returns to education in more remote regions[75]. Although as with gender, we speculate that this can be a 'correction or imbalance' factor at work. The impact of distance has similarities to a gravity model, as well as other spatial or geographic models, whereby distance from major urban centres and/or ports impacts upon factors such as productivity, trade and

[75] We have also implemented separate regressions (panel estimates with random effects) by different regions with the inclusion of year fixed dummies. We obtained statistically significant results for four regions: Liaoning, Jiangsu, Henan and Hubei. The coefficients for these regions are 2%, 3%, 3% and 4% for Liaoning, Jiangsu, Henan and Hubei respectively. We do not report results here.

even housing (Buch, *et al.*, 2004; Roy and Thill, 2004; Rice and Venables, 2003; Redding and Venables, 2004). Our results suggest that the marginal productivity of labour and productivity is higher in Beijing and specifically Shanghai than other areas. The impact of large centres of population on productivity, e.g., is in part because of economies of scale and closeness to market. Closeness to ports and hence export markets is a further important geographical rational for the siting of firms, in particular multinational firms. Multinational firms tend to be more productive than non-multinationals and can by a demonstration effect impact upon the productivity of both supplier and competitor firms. The extent of this impact deteriorates with distance. Thus these results are consistent with the hypothesis that knowledge enhancement, e.g. through process innovation, tends to impact firstly on Shanghai and then become diffused across China at a speed inversely proportional to distance from Shanghai. Shanghai may be the epicentre of this process as this is where there is the greatest density of multinationals. This is consistent, as we have said, with recent work in economic geography.

Chapter 5 The Influence of Socio-Economic Status and Body Mass Index on Health

5.1. Introduction

The interest in the relation between the components of socio-economic status (SES) and self-reported health (SRH) has been renewed within the recent years. Researchers have consistently demonstrated that people with higher socio-economic status have better health, no matter which measures of SES are included or how health is measured (Smith, 1999; 2004). Socio-economic status is often measured by a person's income, educational level, occupation, marital status, or a combination of these factors. Within them, educational attainment, occupational status and income are the three dominant components of SES, and their predictive influences on health have been found by researchers in the United States, the UK, and a number of other European countries. As to Asia, some similar research has been done in Thailand, Philippines and China, but most of the research on China has focused on elderly people only (Liang, *et al.*, 2000; Yi, *et al*., 2002; Zimmer and Kwong, 2004).

The relationship between SES and health has been heavily researched in economics and the other social sciences. The health investment model developed by Grossman (1972) offers an important explanation that *health can be viewed as a durable capital stock that yields an output of healthy time. Individuals inherit an initial stock of health that* *depreciates with age and can be increased by investment.* Grossman and Kaestner (1997) presented an overview on the relation between education and health, and their basic hypothesis is that education can affect a person's knowledge and information on how to promote health. However, there exist plausible reasons why education related differences in health can not be interpreted causally as discussed in Chapter 3. We specifically focus on the hypothesis suggesting that the relationship between education and health may not be causal due to unobservables. Importantly, similar concerns relate to the relationship between income and health. Thus we pursue an approach, using panel data techniques to allow for the unobservables.

Another important aspect of the research in this chapter is the concept of a healthy weight range, which is based on a measurement known as the Body Mass Index (BMI). It is an acceptable proxy for thinness and fatness, and a useful measure of determining obesity and the increased risk of an unhealthy life. The BMI is calculated as weight in kilograms over height in meters squared (weight (kg)/height $(m)^2$). The World Health Organization (WHO) has devised a classification where by persons with BMIs below the range 19-25 are considered underweight, those with BMIs above this range are considered overweight or "at risk", and those with BMIs greater than or equal to 30 are considered obese (Bell, *et al.*, 2002). In addition, recent studies have indicated that the BMI is age and gender dependent when used as an indicator of body fatness (see, e.g., Gallagher, *et al.*, 1996; Burkhauser and Cawley; 2008). However, there are no clear cut-off points after adjusting for age and gender in the literature.

After the start of social economic reform in 1978 in China, there has been increased attention on how to improve the awareness of health problems, and how these problems are distributed across people with different personal and social characteristics. In this chapter, we use the panel data of the Physical Examination in China Health and Nutrition Survey (CHNS) covering the years of 1991, 1993, 1997 and 2000. The main aim of this chapter is to determine how a variety of socio-economic factors influence SRH, and BMI, as an alternative measure of health which is more related to health behaviors. Given the use of a categorical measure of SRH, we estimate a (random-effects) ordered probit model. For comparison and testing purposes we also estimate a fixed-effects ordered logit model using the estimator recently developed by Ferrer-i-Carbonell and Frijters (2004). We therefore attempt to make use of the panel element of the CHNS to particularly deal with the potential problem of unobserved heterogeneity. In addition, we attempt to develop a little further the issue on how to define the BMI with respect to the adjustment of gender and age. Based on our 'new' measure of BMI, we further examine its 'healthy range' based on self-reported measures of health and estimate its impact on SRH together with SES.

This chapter is organized as follows: In the next section, we will provide a comprehensive literature review based on the impact of SES on SRH and BMI. In section 3, we describe the model and data specifications. Preliminary statistics are also provided in this section. In section 4, we outline our econometric framework. Section 5 presents the empirical results. This will be in three stages. Firstly the determinants, especially education and income, of health, followed by the impact of SES on the documented healthy BMI. Finally we develop a little further the issue on how to define the BMI and estimate its impact on SRH. We conclude this chapter in the final section.

5.2. Literature Review

5.2.1. SES and Health

The association between SES and health has been well-documented in the literature, and one of the most consistent social epidemiological findings is the positive correlation between SES and health. In this section, we describe the linkage between a set of economic and demographic variables that are included in nearly every analysis of health status. These variables include education, income, occupation, age, gender and marital status which are correlated with health in one context or another as documented in the literature. For example, Groot and Maassen van de Brink (2007) reported that the health status declined with age and generally older people experience more risk of diseases. The health status of women tends to be less than that of men, although it has been argued that gender differences have more of a mixed impact (Matthews, *et al*., 1998, 1999). For example, gender differences can at least partly be ascribed to age differences between men and women (Groot and Maassen van de Brink, 2007). In addition, married people are healthier than their unmarried peers, and the linkage between marriage/cohabitation and health is much stronger for men than women (Janzen and Muhajarine, 2003; Fuchs, 2004).

Some occupations are less healthy than others. The job demands and conditions of workers in different positions in the occupational hierarchy may differentially affect their health. Studies have clearly shown that job conditions such as high work strain, poor work support, low control in the work place, job insecurity, repetitive work, and exposure to uncomfortable working conditions are associated with elevated risks of health problems, including a general worsening of self-reported health and an increase in specific illnesses such as coronary heart disease, cardiovascular disease and mortality (e.g., Bosma, *et al*., 1997; Borg and Kristensen, 2000). For example, Fuchs (2004) concluded that work related mortality in construction and transportation are triple the rate for all workers. In addition, researchers have also argued that some of the variation in healthy life style choice appears to be related to the observed characteristics of individuals in certain occupations (Contoyannis and Jones, 2004). Winkleby, *et al.* (1992) reported that within occupations, men and women white-collar executives and managers exhibited lower levels of smoking. Among men, executives and managers showed the lowest mean levels of blood pressure. However arguably this reflects the impact of knowledge on behavior as much as the specific nature of the occupation.

Note that the causal mechanisms underlying this relationship between SES and health are complex and controversial. Here we review the recent studies that have attempted to establish the causality, and we mostly focus on the linkage between education, income and health.

5.2.1.1. The Causality between Education and Health

The education-health relationship has received more attention in this area. Different formal educational years or different levels of education have been one of the main determinants of different health status. A negative linkage from education to health is very unlikely (Fuchs, 2004). Just as education is believed to enhance market productivity, education may also enhance health productivity as argued by Grossman (1972). To put it differently, the more educated obtain a large health output from given amounts of inputs. A number of studies have looked at the direct effects of education on health, and almost all of them have reported that higher education strongly contributes to better health (Grossman, 1972; Hay, 1988; Gilleskie and Harrison, 1998; Arendt, 2005; Groot and Maassen van de Brink, 2007). Monden, *et al*. (2003) further argued that not only own education but a partner's education influences health status. In addition, education may also enhance knowledge and information on how to live a healthy life, which is closely linked to healthy behavior. A positive relation between education and healthy behavior is well-documented in the literature. For example, Grossman and Kaestner (1997) presented a comprehensive overview of studies on this relation and reported that higher educated people are less likely to smoke, more likely to exercise more, wear seatbelts more often, and participate in screening programs for breast cancer and cervix cancer. More recently, Al Riyami, *et al*. (2004) reported that women's education had the strongest effect on their use of contraception in Oman excluding traditional factors. There might be also non-behavior reasons for this linkage, related to psychological distress as described by Mulatu and Schooler (2002). Generally, lower SES people tend to be more vulnerable than higher SES ones to psychological distress (e.g., depression), and that such psychological disturbances account, at least in part, for the disproportionate burden of disease among them (Gallo and Matthews, 1999).

There exists the question of whether it is a unidirectional causal relation from education to health. We focus on hypotheses suggesting that the relationship between education and health may be due to unobservables, it thus gives rise to a standard missing variables problem. One important potential missing variable is childhood health, or health endowment shaped prior to educational attainment, which has a long history in the empirical work. Grossman (1975) corrected for this ‘omitted’ variable between education and health by including proxies, for example, using family background and test scores as control variables. He found large and significant direct impact of education on both health and mortality. Another approach is the use of household fixed effects. For example, Behrman and Wolfe (1989) used education of the respondents’ sisters as a control for endowment, and got an estimate of the true education effect as the difference between the coefficients on own education and sisters’ education as suggested by Chamberlain (1980). The result suggested that controlling for household unobservables does not alter education effects considerably. A second explanation for a correlation between an unobserved component and education states that individuals with higher preferences for the future are more likely to engage in activities with current costs and future benefits such as education and health investments. Grossman and Kaestner (1997) discussed studies investigating the time preference hypothesis. They argued that the previous studies tended to deal with health behaviour, notably smoking, and therefore did not deal with the education effects on health per se. They concluded that the existing, sparse evidence does not favour the time preference hypothesis.

A number of studies have dealt with endogeneity of education using instrumental variable techniques. Berger and Leigh (1989) used per capita income and per capita expenditures on education in the state of birth as instruments for education in their study of education effects on health on US data. They found that education effects are slightly reduced, but remain significant, when correcting for endogeneity. Adams (2001) adapted the identification strategy suggested by Angrist and Krueger (1991), using quarter of birth in a US cross-section of individuals from 1992. He used SRH and alternative measures which describe functional limitations. When instrumental variable estimation is applied, education effects increase slightly but are insignificant. Furthermore, F-values on the instruments are just above 1, indicating a problem of weak instruments (Arendt, 2005). Spasojevic (2003) used changes in compulsory schooling in Sweden in the fifties, analyzing education effects on a constructed index of bad health and on an index of healthy BMI range over the period 1981-1991. She found a positive and significant impact of years of education on both measures. Similarly and more recently, Arendt (2005) used Danish school reforms to instrument education effects on health and BMI. He argued that education is related to these two measures in the expected way and the relationships are amplified in magnitude when education is instrumented. However, as the standard errors also increase when using instruments, neither exogeneity nor the null of no effect of education can be rejected for health and BMI.

As it stands, the existing evidence suggests that there is a causal component running from education to health. However, the evidence is still sparse, and many studies suffer either from use of poor instruments, low precision or from a focus on very narrow samples (Arendt, 2005).

5.2.1.2. The Causality between Income and Health

The relationship between income and health is complex, and it is one of the most heavily research topics in economics and other social science. It relates to several aspects including a comparative analysis between countries and over time. Researchers have employed a range of different ways to measure income over time, which can be roughly grouped into income levels over a number of years, income change over time and duration of poverty experience (Benzeval, *et al*., 2000). Most of the studies that include measures of income levels over time and income change found a significant correlation with health outcomes. Greater income is associated with better health through, for example, a better lifestyle or diet, fewer monetary worries, better access to medical services and an improved living environment (Adler, *et al*., 1994; Smith, 1999). Fuchs (2004) argued that this correlation can vary from highly positive to weakly negative, depending on the context, covariates, and level of aggregation. For instance, Hemstrom (2005) reported that the lowest income quintile had better health than those in the second in Sweden, although only for men. In addition, Benzeval and Judge (2001) argued that there is a linkage between low income and poor health, and long-term income is more important for health than current income.

However, similar to education, the main direction of causality between income and health is open to debate. In order to establish the causal effect of income on health, recent studies have attempted to use ‘random variation’ in income, as well as more advanced panel data econometric techniques (Frijters, *et al*., 2005). To explore the effect of income on self-assessed health and chronic health limitations, Ettner (1996) adopted a two-stage IV method, where family income was instrumented by respondents’ wage rate and non-earnings income. The author assumed that these variables only impact on health through their effect on family income, and she found a large positive effect of income on health using all the various measures of health. However, as already noted by Meer, *et al*. (2003), this is a strong assumption because it seems intuitive that both employment and non-earnings income will have a direct effect on health. Meer, *et al*. (2003) used a balanced panel of just over 3000 individuals drawn from four waves of the US Panel Study of Income Dynamics (PSID) to explore the effect of recent changes in wealth on a binary measure of self-reported general health (which was collapsed down from a 5-point ordinal scale). The change in wealth was instrumented by information on the value of inheritance received over the last 5 years, and estimates suggest that increased wealth has a very small positive effect on health. Importantly, as discussed by the authors the validity of inheritance information as an instrument is also open to some debate.

Individual heterogeneity plays an increasing role in explaining variation in health status. Benzeval, *et al*. (2000) used data from the first six waves of the British Household Panel Study to study the relationship between very low income, or poverty and a binary self-reported measure of health. Estimates from binary logit models suggested that recent poverty was a strong predictor of health. However, the lack of controls in the analysis for unobservable heterogeneity, and the lack of exogenous variation in income, indicates that this result should be considered with some caution (Frijters, *et al*., 2005). Some recent studies have started to focus on this issue in the empirical analysis. Contoyannis, *et al*. (2004) used a categorical indicator of self-assessed health, and information on both equivalised and non-equivalised annual household income to explore the causal effect of income on health. They estimated random-effects ordered probit models which allowed for persistence in the observed outcomes due to state dependence and unobserved individual heterogeneity. They explored the issue of state dependence in health by conditioning on previous health status, and they controlled for the problem of initial conditions arising in dynamic panel data model by including individual-specific characteristic averages as suggested by Mundlak (1978). They presented evidence of persistence in self-assessed health explained in part by state dependence which is stronger amongst men than women, and by individual heterogeneity, it accounts for around 30% of the unexplained variation in health.

More recent papers have started to apply the fixed-effects ordinal estimator proposed by Ferrer-i-Carbonell and Frijters (2004) to control for unobserved individual heterogeneity. Frijters, *et al*. (2003) use the unanticipated shift in permanent income for East German households, following the reunification of Germany, as a source of exogenous variation in income. Using the German Socio-Economic Panel (GSOEP) between 1991 and 1999, they fit both random and fixed-effects estimators to the ordinal health measures. Their estimates from the random-effects model suggested that there is a significant positive relationship between income and health. However, after controlling for heterogeneity, there is no evidence for a causal effect of income on health. Based on this discussion, Frijters, *et al*. (2005) further used the GSOEP between 1984 and 2002 and updated the results. They used only the fixed-effects estimator and concluded that there was evidence of a significant positive effect of income changes on health satisfaction, but the quantitative size of this effect is small. This is the case with respect to current income and a measure of 'permanent' income. A more recent study by Jones and Schurer (2007), used the same methodology with 22 waves of the GSOEP (1984-2005), and concluded that adjusting for unobserved heterogeneity accounted for the relationship between income and very good health, but not between income and poorer health status.

5.2.2. SES and BMI

The BMI is an intermediate health indicator and it is a measure of heaviness that reflects both lean and fat issues. The BMI is highly correlated with body fat, and obesity is defined as an excess accumulation of body fat. The BMI is generally calculated as weight in kilograms over height in meters squared (weight(kg)/height(m)^2). As previously mentioned, the WHO has devised a classification where by persons with BMIs below the range 19-25 are considered underweight, those with BMIs above this range are considered overweight or 'at risk[76]' and those with BMIs greater than or equal to 30 are considered obese. These WHO BMI classifications of overweight and obesity are intended for international use. However, as discussed in the chapter 3, the BMI is age and

[76] They reflect risk for type 2 diabetes and cardiovascular disease, which are rapidly becoming major causes of death in adults in all populations (WHO expert consultation, 2004).

gender dependent when used as an indicator of body fatness (see, e.g., Gallagher, *et al*., 1996; Burkhauser and Cawley; 2008), and the BMI cut-off points appear likely to differ across ethic groups (see, e.g. Moon, *et al.*, 2002).

From the economic point of view, some researchers have taken BMI as an element of a life style which is closely related to health behaviours (Contoyannis and Jones, 2004). It is influenced by various factors including age, gender, race, SES and behaviour aspects of life style, such as diet, physical activity and smoking (Reddy, 1998). For example, black people often display higher BMIs than other races (e.g. Mokdad, *et al*., 1999; Lavie, *et al.*, 2004; Goodman, *et al*., 2005), and black females may be particularly at risk of becoming overweight (e.g. Crawford, 2001). Body mass and the prevalence of obesity have been shown to be inversely associated with SES in the United States and other industrialized countries (Sobal and Stunkard, 1989; Jeffery and French, 1996; Montgomery, *et al*., 1998; Wardle, *et al*., 2004). However, for developing countries, the positive association between SES and BMI has been observed in many studies. For example, de Vasconcellos (1994), using Brazilian surveys, reported a higher probability of having thin individuals at lower income levels. Delpeuch, *et al*. (1994) presented a large prevalence of low BMI in rural areas and high BMI in urban areas of the Congo. More recently, Reddy (1998) reported similar results for the India.

For the relevance of this chapter, we are not going to review the impact of behaviour aspects on BMI, however, it is necessary to mention that most of the studies reviewed here considered these aspects together with SES on BMI.

5.3. Model and Data Specification

5.3.1. Theoretical Foundation and Model

Building on the human capital theory, Grossman (1972) provides a formal model to analyze health capital – the Grossman Model. Health is to be viewed as a dimension of human capital, similar but not identical to education. Health increases income through adding healthy working days while education through improving productivity (Zhao, 2005). In this model, health depreciates over time but is maintained by health investment depending on consumed goods and activities that affect health (Arendt, 2005). Gross investment is produced by a household production function, which is affected by the efficiency or productivity of a given consumer as reflected by his or her personal characteristics. Specifically, efficiency is defined as the amount of health obtained from a given amount of health input, for example, years of education completed plays a large role in this context (Grossman, 2000).

Our empirical study will test these theoretical implications, and our models should be viewed as reduced form specifications as they do not include objects of choice, such as medical care, or other health inputs, such as lifestyle. We hypothesize that a person's health is linked to their socio-economic characteristics apart from their personal characteristics, such as education, income, etc. The rationale is found in the literature. An individual will attempt to maximize health capital subject to time, information constraints (the stock of knowledge) and financial status. To be consistent with the literature, we also control other factors such as gender, marriage status, occupation and region, which potentially model life style. The specification is as follows:

$$\text{Health} = \beta_0 + \beta_1\text{age} + \beta_2\text{age}^2 + \beta_3\text{gender} + \beta_4\text{education} + \beta_5\text{income} +$$
$$B_6\text{married} + \beta_7\text{occupation} + \beta_8\text{urban} + \beta_9\text{region} + \varepsilon \qquad (5.1)$$

5.3.2. Data Specification

The dataset we use comes from the Physical Examination in China Health and Nutrition Survey which is a four-year panel data survey including 1991, 1993, 1997 and 2000. This dataset was collected mainly by the Carolina Population Centre and it is designed to examine the effects of the health, nutrition, and family planning policies and programs implemented by national and local governments. It provides a valuable national sample for researchers in health and nutrition fields, so it is suitable for our purposes. CHNS utilizes a multistage, random cluster-sampling scheme. The sample households were randomly drawn from eight provinces including Liaoning/Heilongjiang, Shandong, Jiangsu, Henan, Hubei, Hunan, Guangxi, and Guizhou[77], and in each province, both rural and urban residents are sampled. We restrict our sample to those between 15 and 75. After also excluding observations with less than full information, it provided 27882 observations (14285 males; 13597 females) in the four years altogether. All the variables are defined in Table 5-1, and Table 5-2 shows sample means.

[77] Liaoning was replaced by Heilongjiang in 1997 and both Liaoning and Heilongjiang were included in 2000.

Table 5-1 Data Description

Variables	Data Description
Education (Educ)	A continuous value from 0 to 18
Gender	A dummy variable: Males (‘1’) & Females (‘0’)
Age	A continuous value, restricted from 15 to 75. Age and age-squared are used in the main regressions.
Marital Status	A dummy variable: Married (‘1’) & Non-married (‘0’). Non-married includes never married, divorced, widowed and separated.
Living Area	A dummy variable: Urban (‘1’) & Rural (‘0’)
Household Income (ladhinc)	Total deflated (by 1989 price index) annual household income, log value is used in the regression.
Occupation	Dummy variables. Seven occupations have been chosen based on the sample size, which include: Professionals[1], Administrators, Office Staff, Farmers, Skilled-workers, Drivers and Service Workers.
Regions	Dummy variables. Nine provinces include Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi and Guizhou.
Illness	A dummy variable reflecting any difficulty in carrying out daily activities and work due to illness: Yes (‘1’) & No (‘0’).
BMI	Body Mass Index (weight (kg)/height (m)²) – Healthy range is identified with BMIs of 19-25[2].
Self-reported Health	Ordinal scales: Excellent (‘3’), Good (‘2’), Fair (‘1’), and Poor (‘0’)[3].

Notes: [1] Professionals include senior professional/technical personnel (doctor, professor, etc.) and professional/technical personnel (editor, photographer, etc.); Administrators include executive/manager, factory manager, government official, section chief, director, administrative cadre; and Service Workers include housekeeper, cook, waiter, doorkeeper, barber/beautician, counter sales, launderer, and childcare. The remaining occupation is the base category in the regression which includes such as non-technical/ non-skilled worker (laborer), homemaker-with no other work, student and others. [2] The figures come from the World Health Organization. [3] We have reversed the scales of the health measure in the CHNS survey to emphasize that higher numbers correspond to better health.

As mentioned before, educational attainment, occupational status and income are three dominant components of SES. In the survey, completed years of formal education were measured by primary school (1-6 years), lower middle school (1-3 years), upper middle school (1-3 years), middle technical or vocational school (1-2 years) and college/ university (1-6 years or more). We aggregate these discrete values based on China’s education system to obtain continuous values of the formal education years[78]. It thus takes a value from 0 to 18 with an average value of 6.6 years (7.51 for males and 5.66 females) for all four years together. From Table 5-1, we can see the information on different types of occupations in the survey. Because the number of the observations in some types of occupations is relatively small, we select seven occupations which are professionals, administrators, office staff, farmers, skilled-workers, drivers and service workers. Income is measured as deflated[79] total annual household income (ladhinc). It is the sum of household incomes from all sections including income from wages, home gardening, household farms or farming collectives, raising livestock/poultry, collective and household fishing, household business, welfare subsidies or ration coupons, housing subsidies and other sources of income. This variable is transformed to natural logarithms to allow for concavity of the health-income relationship (see e.g. Frijters, *et al*., 2003; Contoyannis, *et al*., 2004). Table 5-2 shows that the average value for the logarithm of this variable is 8.36 for whole sample. We also include data on marital status, and being married is the included category for estimates.

[78] In China, basic formal education includes primary education (normally six years) and secondary education. Secondary education is divided into academic secondary education (normally three years of lower and three years of upper middle school) and specialized/ vocational/technical secondary education, i.e., after graduated from the lower middle school, one can apply for upper middle school or middle technical/vocational school.

[79] According to CHNS, there is no published absolute consumer price index for China that provides a way to compare provinces or urban and rural areas. Rather the State Statistical Bureau publishes annually a consumer price index ratio that shows for urban and rural areas in each province the shift in the cost of living within that geographic area. The CHNS urban and rural price data are used to create a ratio of urban and rural costs for elements of the consumer goods basket, and they create their own costs (yuan) of the consumer basket for each time period for urban and rural areas in each province in the CHNS. Their deflator is based on 1988 prices.

Table 5-2 Sample Means

Variables	Whole Sample		Males		Females	
Self-reported Health	1.84		1.88		1.80	
Males (%)	0.51		-		-	
Education	6.60		7.51		5.66	
Age	40.24		40.62		39.84	
Married (%)	0.80		0.80		0.80	
Urban (%)	0.29		0.29		0.29	
Ladhinc	8.36		8.36		8.36	
Professionals (%)	0.05		0.06		0.05	
Administrators (%)	0.04		0.07		0.02	
Office Staff (%)	0.03		0.04		0.03	
Farmers (%)	0.57		0.51		0.62	
Skilled-workers (%)	0.07		0.09		0.05	
Drivers (%)	0.01		0.03		0.00	
Service workers (%)	0.07		0.05		0.08	
Liaoning (%)	0.07		0.07		0.07	
Heilongj (%)	0.05		0.06		0.05	
Jiangsu (%)	0.13		0.13		0.13	
Shandong (%)	0.11		0.11		0.10	
Henan (%)	0.12		0.12		0.12	
Hubei (%)	0.13		0.13		0.13	
Hunan (%)	0.11		0.12		0.11	
Guangxi (%)	0.14		0.14		0.14	
Guizhou (%)	0.14		0.14		0.14	
No difficulty in carrying out daily activities and work† (%)	95.68		95.94		95.41	
BMIs in healthy range (19-25) (%)		0.72		0.73		0.70
BMIs underweight (<19) (%)		0.13		0.13		0.13
BMIs overweight (>25) (%)		0.15		0.14		0.17
No. of the Observations	27882	26366	14285	13356	13597	13010

Source: Physical Examination in China health and Nutrition Survey, various years.
Notes: † No. of the observations for this variable are slightly less than others due to missing values. It is only used for the test of the validity and reliability of self-reported health in the empirical section.

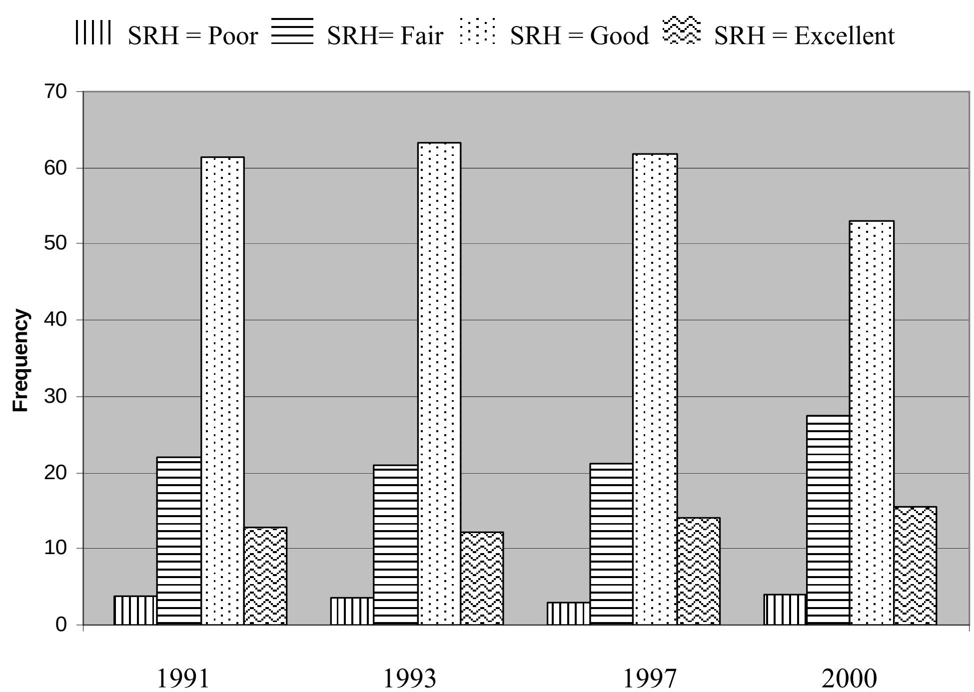
The main measure of health that we use is self-reported health status which is defined by a response to 'how would you describe your health compared to that of other people of your age?' The responses to this question take the ordered scale: poor, fair, good and excellent. SRH has been used widely in previous studies of the relationship between health and SES (see e.g. Ettner, 1996; Smith, 1999; Adams, *et al*., 2003; Frijters, *et al*., 2003; 2005). It is a simple subjective measure of health that provides an ordinal ranking of perceived health status, however, there is evidence to suggest that it is a powerful predictor of subsequent mortality (see e.g. Idler and Kasl, 1995; Idler and Benyamini, 1997) and its predictive power does not appear to vary across socioeconomic groups (see e.g. Burstrom and Fredlund, 2001). SRH have also been shown to be a good predictor of future medical care usage (see e.g. van Doorslaer, *et al*., 2000; 2002).

In addition, note that the BMI is more closely related to health behavior than SRH (Arendt, 2005), we also report some results using the documented BMI healthy range (19-25) as a robustness check. The BMI healthy range is calculated as follows:

$$
\begin{aligned}
&BMI_u' = BMIs < 19 \\
&BMI_o' = BMIs > 25 \\
&BMI_u = BMI_u'*(19-BMIs) \qquad (5.2)\\
&BMI_o = BMI_o'*(BMIs-25) \\
&\text{Healthy } BMI = BMI_u + BMI_o
\end{aligned}
$$

In equation (5.2), we assume that BMI_u measures the distance of the individuals' BMI values from 19 (if their BMIs are less than 19, underweight), and BMI_o measures the distance of the individuals' BMIs from 25 (if their BMIs are more than 25, overweight). The Healthy BMI indicates the greater the distance from either side the less healthy they are. As expected, the BMI_u and BMI_o are negatively correlated with our SRH measure[80]. Figure 5-1 describes the distribution of SRH across all four years. The distributions show that the majority of observations reflect good health, and there is a trend for the distribution of health to become worse, specifically for the year 2000. Compared to the former three years, this can be seen by the increase in the proportions of observations in the poor and fair categories, while there is a decrease in the proportion reporting good health. Figure 5-2 describes the average health status across all four years and it shows the same trend as indicated in Figure 5-1[81].

Figure 5-1 Self-reported Health Status by Year

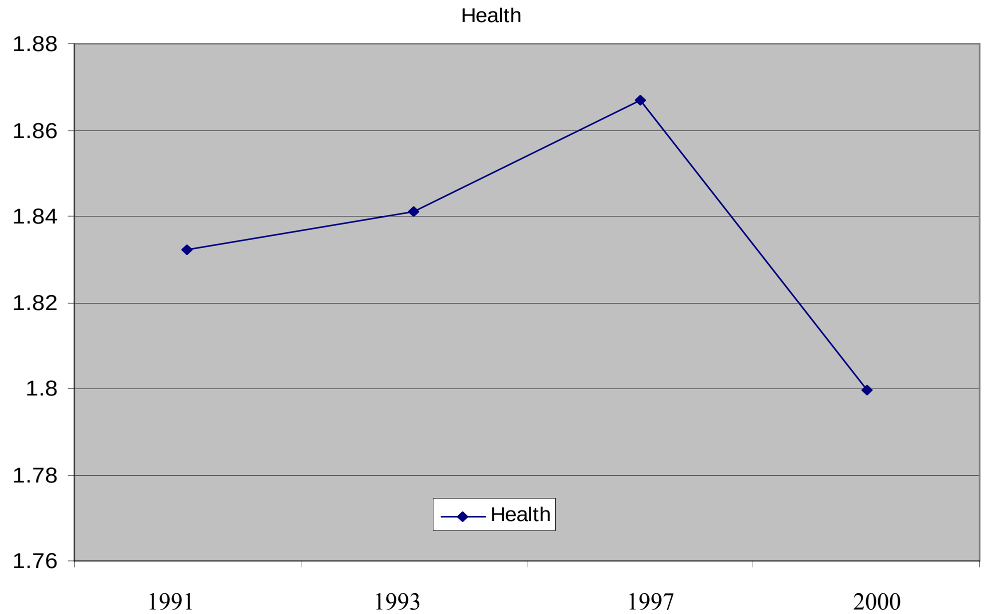


Source: Physical Examination in China health and Nutrition Survey, various years.

[80] The negative correlation between BMI_o and self-reported health is not strongly significant. This is not surprising in developing countries, such as China, since ill-nutrition is still a major cause for poor health.

[81] To rule out the possible health-related attrition in the data, we redo Figure 5-2 using balanced data with 9216 observations, and we get a very similar figure.

Figure 5-2 Average Status of Self-reported Health by Year



Source: Physical Examination in China health and Nutrition Survey, various years.

Table 5-3 gives descriptive statistics on self-reported health across gender, urban/rural and province. 11.4% of urban residents reported to have excellent health, and this percentage is relatively lower compared to rural ones (14.3%). Compared to men, lower percentage of women reported to have excellent and good health. People in Guangxi and Guizhou (western provinces) have the lowest percentages of excellent health status. They are 4.4% and 8.3%, respectively. People in Liaoning, Heilongjiang (northeastern provinces), Jiangsu and Shandong (coastal provinces) have the highest percentages of excellent health. The difference is striking. However, if we combine the two categories of excellent and good health, the gap between western provinces and other provinces becomes smaller. In all provinces, less than 5% people report to have poor health. However, it is important to keep in mind that this kind of purely descriptive analysis that exploits solely the cross-sectional variation in the data is not able to reveal anything about the underlying causal relationship between these variables and health.

Table 5-3 Self-Reported Health by Genders, Urban/Rural and Provinces

		Liaoning	Heilongjiang	Jiangsu	Shandong	Henan	Hubei	Hunan	Guangxi	Guizhou
Poor	Freq.	81	38	104	53	155	144	96	157	160
	%	3.9%	2.6%	2.9%	1.8%	4.7%	4.0%	3.1%	4.0%	4.0%
Fair	Freq.	417	261	709	486	849	858	715	1179	855
	%	20.3%	18.0%	20.0%	16.5%	25.7%	24.1%	22.9%	30.2%	21.5%
Good	Freq.	1221	736	2058	1708	1810	2284	1961	2392	2636
	%	59.5%	50.6%	58.0%	57.9%	54.7%	64.1%	62.7%	61.4%	66.2%
Excellent	Freq.	332	419	678	702	493	279	357	170	329
	%	16.2%	28.8%	19.1%	23.8%	14.9%	7.8%	11.4%	4.4%	8.3%
Total	Freq.	2051	1454	3549	2949	3307	3565	3129	3898	3980
	%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 5-3 continued

			Whole Sample	Liaoning	Heilongjiang	Jiangsu	Shandong	Henan	Hubei	Hunan	Guangxi	Guizhou
Female	Poor	Freq.	512	39	20	54	31	84	73	50	90	71
		%	3.8%	3.9%	3.0%	3.1%	2.2%	5.1%	4.1%	3.5%	4.7%	3.6%
	Fair	Freq.	3352	227	139	397	260	442	449	355	639	444
		%	24.7%	22.7%	20.8%	22.6%	18.4%	27.1%	25.3%	24.5%	33.1%	22.5%
	Good	Freq.	8130	596	332	1009	808	884	1141	911	1132	1317
		%	59.8%	59.5%	49.7%	57.4%	57.3%	54.1%	64.4%	62.9%	58.6%	66.9%
	Excellent	Freq.	1603	140	177	299	312	224	110	132	71	138
		%	11.8%	14.0%	26.5%	17.0%	22.1%	13.7%	6.2%	9.1%	3.7%	7.0%
	Total	Freq.	13597	1002	668	1759	1411	1634	1773	1448	1932	1970
		%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Male	Poor	Freq.	476	42	18	50	22	71	71	46	67	89
		%	3.3%	4.0%	2.3%	2.8%	1.4%	4.2%	4.0%	2.7%	3.4%	4.4%
	Fair	Freq.	2977	190	122	312	226	407	409	360	540	411
		%	20.8%	18.1%	15.5%	17.4%	14.7%	24.3%	22.8%	21.4%	27.5%	20.4%
	Good	Freq.	8676	625	404	1049	900	926	1143	1050	1260	1319
		%	60.7%	59.6%	51.4%	58.6%	58.5%	55.3%	63.8%	62.5%	64.1%	65.6%
	Excellent	Freq.	2156	192	242	379	390	269	169	225	99	191
		%	15.1%	18.3%	30.8%	21.2%	25.4%	16.1%	9.4%	13.4%	5.0%	9.5%
	Total	Freq.	14285	1049	786	1790	1538	1673	1792	1681	1966	2010
		%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 5-3 continued

			Whole Sample	Liaoning	Heilongjiang	Jiangsu	Shandong	Henan	Hubei	Hunan	Guangxi	Guizhou
Rural	Poor	Freq.	702	54	12	74	39	124	90	64	113	132
		%	3.5%	4.2%	1.2%	2.9%	1.7%	5.3%	3.6%	2.9%	4.2%	4.6%
	Fair	Freq.	4233	261	112	431	367	556	548	510	809	639
		%	21.4%	20.1%	11.1%	16.6%	16.1%	23.6%	22.0%	23.0%	30.1%	22.5%
	Good	Freq.	12008	777	544	1511	1336	1297	1655	1419	1655	1814
		%	60.7%	59.8%	53.7%	58.3%	58.7%	55.0%	66.6%	64.0%	61.5%	63.8%
	Excellent	Freq.	2833	208	345	574	535	382	193	224	112	260
		%	14.3%	16.0%	34.1%	22.2%	23.5%	16.2%	7.8%	10.1%	4.2%	9.1%
	Total	Freq.	19776	1300	1013	2590	2277	2359	2486	2217	2689	2845
		%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
Urban	Poor	Freq.	286	27	26	30	14	31	54	32	44	28
		%	3.5%	3.6%	5.9%	3.1%	2.1%	3.3%	5.0%	3.5%	3.6%	2.5%
	Fair	Freq.	2096	156	149	278	119	293	310	205	370	216
		%	25.9%	20.8%	33.8%	29.0%	17.7%	30.9%	28.7%	22.5%	30.6%	19.0%
	Good	Freq.	4798	444	192	547	372	513	629	542	737	822
		%	59.2%	59.1%	43.5%	57.0%	55.4%	54.1%	58.3%	59.4%	61.0%	72.4%
	Excellent	Freq.	926	124	74	104	167	111	86	133	58	69
		%	11.4%	16.5%	16.8%	10.8%	24.9%	11.7%	8.0%	14.6%	4.8%	6.1%
	Total	Freq.	8106	751	441	959	672	948	1079	912	1209	1135
		%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Source: Physical Examination in China health and Nutrition Survey, various years.

5.4. Econometric Framework

Considering cross-sectional data, the ordered probit model arises when considering an independent sample of data (y_i, x_i) where the dependent variable y_i has M possible outcomes (0, 1, 2, …, M) with a 'natural' ordering (Wooldridge, 2002). The ordered probit model for y_i (conditional on explanatory variables x_i) can be derived from a latent variable model. Assume that a latent variable y^* is determined by

$$y_i^* = x_i^{'}\beta + u \quad for\ i=1, 2, \ldots, n,\ u \mid \mathrm{x} \sim \mathrm{N}\ (0,1) \qquad (5.3)$$

This applies to our measure of self-reported health, which has categorical outcomes of poor, fair, good, and excellent. The latent variable, y_i^*, is assumed to be a function of a vector of SES x_i. Define the following observability criterion $y_i = \lambda$ if

$$\lambda_{k-1} \leq y^* \leq \lambda_k \quad for\ k=1, 2, \ldots K. \qquad (5.4)$$

Let $\lambda_0 < \lambda_1 < \ldots < \lambda_M$ be known cut points (or threshold parameters), and $\lambda_0 = -\infty$ and $\lambda_K = \infty$. The conditional probability of observing $y_i = k$ is

$$\mathrm{P}(y_i = k \mid x_i) = \mathrm{P}(\lambda_{k-1} \leq y^* \leq \lambda_k \mid x_i) = \mathrm{P}(\lambda_{k-1} \leq x_i^{'}\beta + u \leq \lambda_k \mid x_i) \qquad (5.5)$$

Given the standard normal assumption for u, rearranging terms gives

$$P(y_i = k \mid x_i) = P(\lambda_{k-1} - x_i^{'}\beta \leq \ u \leq \lambda_k - x_i^{'}\beta \mid x_i) = \Phi(\lambda_k - x_i^{'}\beta) - \Phi(\lambda_{k-1} - x_i^{'}\beta) \qquad (5.6)$$

where Φ(.) is the cumulative normal distribution function. Other distribution functions can be used in place of Φ. Replacing Φ with the logit function (symmetrically assuming a logistic distribution), Λ, gives the ordered logit model.

Both these models have been often used with cross-sectional data in analysis about health satisfaction and well-being (D'Addio, *et al*., 2003). The ordered probit/logit model has also been used in panel studies. In that case unobserved heterogeneity has been dealt within the random effects model. The fixed effects model has been rarely followed owing to the lack of suitable econometric approaches. However, some authors have adopted it by transforming the ordinal variable into a binary one that takes the value of one above or under a specific threshold. To estimate the impact of SES on health, we use pooled ordered probit, the random-effects ordered probit model and the fixed-effects ordered logit estimator recently proposed by Ferrer-i-Carbonell and Frijters (2004).

Our dependant variable, health, $H \in \{0, 1, 2, 3\}$ is an ordinal indicator of health as evaluated by individuals. This measure is available for a set of individuals indexed by $i=1, 2, \ldots,n,$ each observed over a given time-period indexed by $t = 1, 2 ,\ldots, T.$ More precisely, we observe a four-year panel data including 1991, 1993, 1997 and 2000. For each year in which H_{it} is observed, we also observed a (row) vector x_{it} contain a set of covariates describing the characteristics and situation of individual i in year t as mentioned in section 5.3.1.

5.4.1. Random Effects Ordered Probit

As our baseline model we begin with pooled ordered probit. Further we fit the ordered probit model with individual random-effects:

$$H^{*}_{i,t} = x'_{i,t}\beta + v_i + \varepsilon_{it}$$

$$H_{i,t} = k \Leftrightarrow H^{*}_{i,t} \in [\lambda_k, \lambda_{k-1}) \qquad (5.7)$$

where H^{*}_{it} is latent health; H_{it} is observed health; λ_k is the kth cut-off point (increasing in k) for categories; x_{it} are observable individual characteristics; v_i is an individual normally-distributed random characteristic that is orthogonal to x with unknown variance; and finally ε_{it} is a time-varying error-term, normally-distributed, orthogonal to all x . As in the binary choice model, the underlying variance $\sigma^2 = \sigma_v^2 + \sigma_\varepsilon^2$ is not identified. We adopt the normalization $\sigma_\varepsilon^2 = 1$, thus the overall error variance equals (σ_v^2+1).

The associated log-likelihood function for this model is well-established and can be generalized from the arguments made by Butler and Moffit (1982). Individual heterogeneity is unobserved, and therefore to obtain the unconditional log-likelihood we need to integrate the conditional log-likelihood. The integration is done using the

Gauss-Hermite quadrature with 12 evaluation points. Frechette (2001) provides a derivation of the likelihood function for this model and a further discussion of the Gauss-Hermite quadrature estimation.

5.4.2. Conditional Fixed Effects Ordered Logit

It is very likely that there are some unobservable individual traits and characteristics that are related to current health (Frijters, *et al.,* 2003). For example, parental background plays an important role in determining both adult health and current income (Case, *et al*., 2002). Moreover, the relationship between marriage and health may involve selection effects in that partners select each other on the basis of health. Similarly, healthier students may be more efficient producers of additions to the stock of knowledge through formal schooling, and healthier individuals may be more likely to have a job. In those cases, a spurious correlation between marriage, education, work and income and those unobserved characteristics may arise and thereby bias the estimation coefficients. While the random-effects ordered probit can to a certain extent indicate the direction of the effects of some determinants of health, the above-mentioned spurious correlation is still likely to be present (D'Addio, *et al*., 2003). Therefore, a fixed-effects approach seems to be more appropriate.

The following fixed-effects ordered logit model was proposed by Ferrer-i-Carbonell and Frijters (2004):

$$H_{it}^{*} = x_{it}^{'}\beta + f_i + \varepsilon_{it}$$

$$H_{it} = k \Leftrightarrow H_{it}^{*} \in [\lambda_k, \lambda_{k-1}) \tag{5.8}$$

where H_{it}^{*} is latent health; H_{it} is observed health; λ_k is the kth cut-off point (increasing in k) for categories; x_{it} are observable time-varying characteristics; f_i is an individual fixed characteristics; and ε_{it} is a time-varying logit-distributed error-term that is orthogonal to all x . This model is an ordered logit model with fixed individual effects and an individual specific threshold.

This model is an extension of the conditional fixed-effects logit model by Chamberlain (1980) for dependant variables with more than two categories which recodes the data crossing over a barrier, that is the same for everyone, k. The fixed-effects estimator proposed by Ferrer-i-Carbonell and Frijters (2004) is the likelihood of observing which of the T satisfactions of the same individual are above k_i, given that there are c of T satisfactions above k_i. $S(k_i, c)$[82], which denotes the set of all possible combinations of $\{H_{i1}, H_{i2}, ..., H_{iT}\}$ such that $\sum_t I(H_{it} > k_i) = c$. Noteworthy is the fact that the model used in our analysis slightly deviates from theirs. In our case, the recording of observations to 'high' or 'low' health is also individual specific, however, it is based on the individuals' average health scores in the panel[83]. In this case, only individuals with

[82] For detailed discussions see Ferrer-i-Carbonell and Frijters (2004).
[83] See also Brenner (2006), Böckerman and Ilmakunnas (2007) and Jones and Schurer (2007). Further, Dr. Ferrer-i-Carbonell has also mentioned that the two versions of this model presented very similar results.

changes in health status over time can be included.

Finally, we add year dummies to these models to capture varying respondent behavior which is induced by institutional and other changes that affect each respondent in the same fashion (Brenner, 2006). Further, one important methodological point concerns the use of the fixed-effects estimators. According to Frijters, *et al*. (2005), one cannot simultaneously include age, time and fixed-effects in the analysis. Note that $\text{age}_{it}\beta_{\text{age}}=\text{age}_{i0}\beta_{\text{age}}+t\beta_{\text{age}}$, the effect of $\text{age}_{i0}\beta_{\text{age}}$ is time-invariant and will therefore be in the individual fixed-effects that drops out, and $t\beta_{\text{age}}$ will be the same for everyone at t and hence be captured by year dummies. Consequently, the year dummies will also capture age effects.

5.5. Empirical Analysis

5.5.1. SRH – (Random-Effects) Ordered Probit

The former two columns of Table 5-4 present the coefficient estimates for the ordered probit models based on pooled and random effects specifications. The estimated coefficients for these two models are extremely similar, although they are not directly comparable due to different scaling of the error variance. However, we can compare the relative effects of pairs of variables across these models. The coefficient of education on self-reported health is significantly positive at the 1% level for both models, which shows that a one-unit increase in education increases the probability of being in better health. This is consistent with the literature as mentioned in section 5.2.

As to the gender differences, our results show that males are more likely to have better health than females. Self-reported health declines with age, which is consistent with intuition, and the rate of decline increases as indicated by the negatively significant coefficient of age-squared. Surprisingly perhaps, marriage does not appear to be significantly associated with better health status. To estimate and further capture the impact of married status on SRH, we construct an interactive variable Married*Young[84] (age less than 45[85]). The coefficient estimates of this interaction term for the ordered probit models based on pooled and random effects specifications are shown in the third and fourth columns of Table 5-4. Married people are less healthy than non-married except for young couples who are healthier. To put it differently, the positive linkage between being married and better health status is only for young couples. The reasons behind this may be rather complex. One possibility is when people get older, they have more responsibilities and pressures from family and society as they have both children and parents to be taken care of, and they may also use their money differently compared to the young couples as they have to save for the future, for example, if they are not in one of the health care schemes, which is to some extent related to the nutrition intake and healthy life style.

There has existed some urban-rural health inequality in China for many decades if not

[84] We have also tried to include another interactive variable Married*Gender, however, it failed to indicate any significant results. Results have not been reported in this chapter.

[85] In late 20s, WHO redefined the age range: The Young (less than 45), The Middle (45-59), Slight Old (60-74), The Old (75-89) and more than 90.

centuries. Health care is more accessible and of better quality in urban areas compared to rural areas as discussed in chapter 2. However, the sign and significance of the 'urban' variable indicates that rural people are more likely to have better health than urban ones, other things being equal. We could possibly explain this depending on environmental factors, such as the amount of physical activities. Individuals living in rural areas usually have more outdoor activities than urban ones based on the rural characteristics. In addition, there is comparatively less pollution and lower traffic volumes in rural areas which are also important factors in determining better health (see e.g., Frumkin, 2002).

The positive linkage between household income and health has been largely observed as mentioned in the section 5.2. Our results are consistent with the literature and show that more household income significantly increases the probability of being in better health. In addition, most occupations have a significant impact on health status. Farmers are less likely to have better health. Except for professionals, five other occupations have more possibility to being healthy than the default case (i.e. the occupations apart from the seven included in the regression). The reasons are possibly related to a combination of factors including features of the work environment, social circumstances outside work, job control across occupational classes and life style factors. Those executives, office staff and skilled workers are so-called 'white collar', they have comparatively better work environment and higher socio-economic position. As to drivers, it is reasonable to say that they have an extremely high degree of control over their jobs.

Region is an important determinant of health. We take Shandong province, one of the developed coastal provinces in the survey, as the reference province. Table 5-4 shows that all regional coefficients are significantly negative except for Heilongjiang which failed to be significant. These results indicate that individuals living in coastal provinces (Shandong and Jiangsu) and the north-eastern province (Liaoning) have the highest possibility of being healthy, however, people living in Guangxi (one of the western provinces) have the lowest possibility of being healthy.

The fifth and sixth columns of Table 5-4 present the coefficient estimates with the inclusion of a dummy variable which reflects any difficulty in carrying out daily activities and work due to illness. Several studies used the ability to perform normal "activities of daily living" (ADL) as another self-reported measure of health, but most of the research has focused on elderly people only (see e.g. Zimmer, *et al.*, 2004; Beydoun and Popkin, 2005), since these ADL-based measures are closely related to physical functioning[86] or disability. Our results show that the coefficient of this illness dummy on self-reported health is significantly negative at the 1% level for both models. We use this as a test for the validity and reliability of self-reported health in reflecting individuals' general health[87]. The magnitude and significance of the other variables are largely as before.

---

[86] Physical functioning tasks are often classified either as "activities of daily living" (Katz, *et al*. 1963), which include those necessary for personal care, or as "instrumental activities of daily living" (Lawton and Brody, 1969), which are necessary to maintain a living environment (Zimmer, *et al.*, 2004). The detailed discussions on how to define these two measures is out of the scope of this section.

[87] Although different health indicators focus on measuring different aspects of health or the health is measured with different kinds of errors in the different indicators, they have similarities in reflecting individuals' general health.

Table 5-4 The Impact of SES on Self-reported Health

Explanatory Var.	Self-reported Health							
	OProbit	RE-Oprob	OProbit†	RE-Oprob†	OProbit‡	RE-Oprob‡	FE-Ologit	FE-Ologit‡
Gender	0.115***	0.126***	0.120***	0.130***	0.114***	0.124***		
	(8.06)	(7.33)	(8.33)	(7.57)	(7.88)	(7.26)		
Educ	0.012***	0.013***	0.011***	0.013***	0.011***	0.012***	0.053**	0.053**
	(4.96)	(4.70)	(4.79)	(4.56)	(4.55)	(4.34)	(2.30)	(2.25)
Age	-0.014***	-0.014***	-0.014***	-0.014***	-0.013***	-0.014***		
	(-3.90)	(-3.46)	(-3.90)	(-3.49)	(-3.81)	(-3.40)		
Agesq	-0.00007*	-0.00009**	-0.00005	-0.00007	-0.00007*	-0.00009*	-0.00055***	-0.00057***
	(-1.91)	(-2.08)	(-1.32)	(-1.54)	(-1.76)	(-1.94)	(-2.59)	(-2.62)
Married	0.012	0.004	-0.057**	-0.068**	0.014	0.007	-0.189**	-0.183**
	(0.54)	(0.16)	(-2.07)	(-2.17)	(0.61)	(0.27)	(-2.06)	(-1.96)
Married*Young			0.088***	0.092***				
(age less than 45)			(4.16)	(3.88)				
Urban	-0.186***	-0.199***	-0.188***	-0.200***	-0.183***	-0.195***		
	(-11.37)	(-10.23)	(-11.49)	(-10.33)	(-11.06)	(-10.09)		
Ladhinc	0.064***	0.063***	0.069***	0.067***	0.062***	0.062***	0.066***	0.069***
	(7.11)	(6.34)	(7.54)	(6.71)	(6.82)	(6.19)	(2.64)	(2.74)
Profes	-0.047	-0.047	-0.041	-0.041	-0.041	-0.040	-0.025	-0.042
	(-1.29)	(-1.14)	(-1.12)	(-0.99)	(-1.13)	(-0.97)	(-0.17)	(-0.28)
Adminis	0.130***	0.131***	0.133***	0.135***	0.134***	0.134***	0.042	0.016
	(3.40)	(3.06)	(3.50)	(3.15)	(3.48)	(3.12)	(0.32)	(0.12)
Staff	0.107***	0.112**	0.106**	0.111**	0.113***	0.117**	0.098	0.087
	(2.60)	(2.44)	(2.57)	(2.42)	(2.71)	(2.55)	(0.74)	(0.64)
Farmers	-0.048**	-0.058**	-0.045**	-0.056**	-0.040*	-0.050**	-0.184**	-0.192**
	(-2.24)	(-2.42)	(-2.12)	(-2.31)	(-1.86)	(-2.07)	(-2.34)	(-2.39)
Skilled	0.094***	0.097***	0.094***	0.097***	0.100***	0.103***	0.127	0.133
	(3.06)	(2.89)	(3.06)	(2.89)	(3.19)	(3.03)	(1.48)	(1.52)
Drivers	0.164***	0.163**	0.154**	0.153**	0.174***	0.173**	0.31	0.326
	(2.66)	(2.36)	(2.50)	(2.22)	(2.79)	(2.50)	(1.53)	(1.60)
Service	0.089***	0.085**	0.089***	0.085**	0.084***	0.081**	-0.035	-0.027
	(2.82)	(2.43)	(2.81)	(2.43)	(2.61)	(2.31)	(-0.35)	(-0.26)

Table 5-4 continued

Explanatory Var.	Self-reported Health							
	OProbit	RE-Oprob	OProbit†	RE-Oprob†	OProbit‡	RE-Oprob‡	FE-Ologit	FE-Ologit‡
Liaoning	-0.299***	-0.326***	-0.301***	-0.329***	-0.297***	-0.321***		
	(-8.98)	(-8.30)	(-9.05)	(-8.37)	(-8.85)	(-8.21)		
Heilongj	0.047	0.054	0.044	0.051	0.057	0.066		
	(1.26)	(1.24)	(1.18)	(1.16)	(1.50)	(1.50)		
Jiangsu	-0.192***	-0.218***	-0.193***	-0.220***	-0.190***	-0.214***		
	(-6.74)	(-6.37)	(-6.78)	(-6.41)	(-6.60)	(-6.28)		
Henan	-0.361***	-0.396***	-0.360***	-0.395***	-0.349***	-0.380***		
	(-12.45)	(-11.35)	(-12.41)	(-11.32)	(-11.84)	(-10.90)		
Hubei	-0.487***	-0.531***	-0.489***	-0.534***	-0.482***	-0.522***		
	(-17.11)	(-15.49)	(-17.19)	(-15.56)	(-16.74)	(-15.26)		
Hunan	-0.372***	-0.408***	-0.373***	-0.409***	-0.364***	-0.397***		
	(-12.68)	(-11.54)	(-12.72)	(-11.57)	(-12.24)	(-11.24)		
Guangxi	-0.620***	-0.675***	-0.621***	-0.676***	-0.615***	-0.665***		
	(-22.05)	(-19.93)	(-22.10)	(-19.98)	(-21.60)	(-19.70)		
Guizhou	-0.396***	-0.439***	-0.393***	-0.436***	-0.401***	-0.440***		
	(-14.13)	(-13.02)	(-14.02)	(-12.93)	(-14.11)	(-13.07)		
y1993	0.026	0.030	0.027	0.031	0.017	0.021	0.062	0.058
	(1.40)	(1.57)	(1.43)	(1.60)	(0.92)	(1.08)	(1.18)	(1.09)
y1997	-0.005	-0.009	-0.002	-0.007	-0.018	-0.022	-0.001	0.004
	(-0.25)	(-0.47)	(-0.12)	(-0.33)	(-0.92)	(-1.11)	(-0.01)	(0.03)
y2000	-0.061***	-0.078***	-0.056***	-0.072***	-0.064***	-0.079***	-0.143	-0.119
	(-2.97)	(-3.57)	(-2.71)	(-3.29)	(-3.09)	(-3.61)	(-0.86)	(-0.69)
Illness					-0.910***	-0.947***		-1.005***
					(-27.20)	(-26.22)		(-11.51)
Pseudo R^2	0.055		0.056		0.069		0.017	0.029
Log-likelihood	-27132.894	-26957.621	-27124.258	-26950.086	-26304.584	-26162.262	-5997.808	-5767.945
No. of the Obs.	27882	27882	27882	27882	27420	27420	16585	16176

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions are estimated by ordered probit (OProbit), random-effects ordered probit (RE-Oprob) and fixed-effects ordered logit (FE-Ologit). † Regressions with the inclusion of the interaction term Married*Young.
‡ Regressions with the inclusion of an illness dummy reflecting any difficulty in carrying out daily activities and work due to illness.

Finally, both the ordered probit and random-effects ordered probit models replicated the downward trends in health status, specifically from the former three years to year 2000.

5.5.2. SRH – Fixed-Effects Ordered Logit

The fixed-effects model allows us to take account of unobservable heterogeneity, such as family background and personality traits, which might affect both education, income and health. Note that years of education completed for a large amount of individuals does not change much during the survey years, this model specifically allows us to establish how changes in income impact on health. The seventh column of Table 5-4 shows the results. When we allow for fixed-effects, the significant positive relationship between household income and health found by the random-effects model still exists, and the magnitude and significance of the coefficients in these two models are very close. Therefore, the random-effects model has not been far away from the mark in the respect of unobservable heterogeneity.

Moreover, some of the other relationships found in the random-effects model are still apparent. A one-unit increase in education significantly increases the probability of being in better health. Health declines with age at an increasing rate[88], and being married appears to be significantly associated with worse health status, which to some extent is consistent with the results of the random-effects as shown in the fourth column

[88] As argued in the section 5.4.2, age effects have been captured by the year dummies, thus we exclude age in our estimate of fixed effects.

of Table 5-4[89]. However, we need to be careful in interpreting these results. The reasons for this negative linkage are clearly related to the *changes* in people's marital status. Specifically, people are getting less healthy when they change from being married to being non-married. Potentially, there are two reasons for this change: 1) a spouse dies, or 2) couples are getting separated or divorced. In both cases, people are getting depressed prior to the change in status. Thus it is not too surprising that this change is linked to negative health outcomes, and it does not necessarily imply that marriage per se reduces health.

As to the impact of occupations, being a farmer is still associated with worse health. However, the positive relationships found between other occupations, including executives, office staff, skilled workers, drivers and service workers, and health are no longer significant[90]. This could be partly due to the fact that a number of individuals do not change their occupations during the survey period, and hence to a large extent these are 'fixed effects'. In addition, it may also imply that those significant correlations found under the random-effects specifications are most possibly due to selection. For example, the impact of individuals being executives was found to be strongly positive in the random-effects model, but low and insignificant in the fixed-effects model. This indicates that it is possible that those executives are relatively healthy themselves, i.e., the healthier are more likely to be executives in the first place, which would explain the positive correlation picked up by the random-effects result.

---

89 We have run the regression with the inclusion of the interaction term Married*Young. The coefficient is positive, however, it failed to be significant. The results have not been reported in this chapter.

90 The coefficient of service workers in the fixed-effects model is negative and insignificant.

When we allow for fixed-effects, the significantly negative relationship between the dummy variable reflecting difficulty in carrying out daily activities and work due to illness and health found by the random-effects model still exists as shown in the last column of Table 5-4. Gender, urban and provinces do not enter the fixed-effects estimation due to their time-invariant characteristics. In addition, year dummies failed to be significant in the fixed-effects model.

5.5.3. The Change of SRH – An Alternative Measure of Health

As mentioned before, a simultaneous relationship may exist between education and health. We focus on the hypotheses suggesting that the relationship between education and health can not be interpreted causally due to unobserved variables. An important potential unobserved variable is inherent health or health endowments shaped prior to educational attainment, and it is the focus of this section. In addition, as argued by Grossman (2000) this unobserved inherent health is also linked to the reverse causality between health and education.

In order to deal with potential simultaneity between education and health, we assume that the unobserved inherent health is individual specific and constant across time. We further make the following critical hypothesis: the change in health[91] is independent of inherent health, i.e. that all people's health (stochastically) deteriorates at the same level

[91] We assume that health answers are ordinally comparable in the former sections of this chapter, i.e. that it is unknown what the relative difference between health answers is but all individuals do share the same interpretation of each possible answer. Here we assume the answers of health question can be interpreted as cardinal, and take it as a robustness check (see e.g., Ferrer-i-Carbonell and Frijters, 2004).

given their educational attainment. We start from the following equation:

$$H_{i,t} = H_{i,0} - \alpha_i t + \beta E_i t \tag{5.9}$$

$$= H_{i,0} - (\alpha_i - \beta E_i)t$$

Where H_{it} represents heath status for individual i at period t. H_{i0} is individual i's inherent health. E_i represents the education attainment for individual i and α_i includes other SES variables. We assume that the individuals stay at this basic level of health until a certain age – could be 20 or 30. At this time health deteriorates (hence $-\alpha_i t$), but the speed of deterioration depends upon the individuals' level of education which we assume to be constant throughout their adult lives. The rate of decline of health is ($\alpha_i - \beta E_i$). We believe this to be consistent with the commonly viewed interpretation of the impact of education on health. Now we assume that this basic level of education is also linked to inherent health:

$$E_{i,t} = \gamma_0 + \gamma_1 H_{i,0} \tag{5.10}$$

Clearly this unobserved H_{i0} affects both education and health, and makes the causal interpretation difficult. However, if we take the change in health from equation (5.9), we obtain:

$$\Delta H_{i,t} = \beta E_i - \alpha_i \tag{5.11}$$

By comparing equation (5.10) and (5.11), we can see that the basic underlying health H_{i0} has an impact on education, however, it is independent of the change in health as we have defined it.

Based on equation (5.11), we analyze whether SES has an impact on changes in health status. In a sense this is even more important to policy makers as the determinants of health change indicate a more direct way on how to improve people's health. There is a four-year panel in the survey including 1991, 1993, 1997 and 2000, so we can estimate the changes from 1993/1991, 1997/1993, and 2000/1997, and this constitutes a clear three-period panel.

The values of Health Change range from 0 to 6[92], where 0 represents the biggest negative change of health status and 6 represents the biggest positive change. We create a dichotomous variable Real Change[93], which takes a value of one if the individuals have a positive change or no change of their health status (the value of Health Change more than or equal to '3'), and zero otherwise. Through this, we can find the SES pattern for people whose health statuses are getting worse. Out of a total of 15649 person-year observations, 73% individuals' health status is getting better or being the same, and 27% is getting worse. The independent variables included in the regression

[92] The original values of health change are from -3 to 3.

[93] Based on the values of Health Change, we started by fitting the (random-effects) ordered probit. The results show almost no significant impact of SES on Health Change. We assume that this may be in part due to the fact that the change of health is not 'ordered'. The construction of 'Real Change' allows us to capture the determinants which make people's health status getting worse. Note that we also tried to capture the determinants which make people's health status getting better, and again, we did not find any significant impact. We did not report those results in this chapter.

are specified in equation (5.1). We use the pooled probit model and further allow for the random-effects. Table 5-5 shows the results, and the likelihood ratio test shows that the panel-level variance component is important.

Table 5-5 The Impact of SES on Real Health Change (Random-Effects) Probit Model

Explanatory Variable	**Real Health Change**	
	Probit	RE-Probit
Gender	-0.002	-0.001
	(-0.07)	(-0.05)
Educ	0.008**	0.008**
	(2.25)	(2.22)
Log(age)[1]	-0.173***	-0.176***
	(-4.08)	(-4.06)
Married	-0.012	-0.013
	(-0.38)	(-0.40)
Urban	-0.040	-0.041
	(-1.49)	(-1.49)
Ladhinc	0.035**	0.035**
	(2.56)	(2.56)
Farmers[2]	-0.061**	-0.063**
	(-2.27)	(-2.27)
Regional Dummies	yes	yes
Year Dummies	yes	yes
Log likelihood	-9043.2706	-9075.4994
Likelihood-ratio		64.46
No. of the Obs.	15649	15649

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions estimated by probit and random-effects probit (RE-Probit). [1] We have first tried to include age and age-squared in the regression to be constant with the main equations, and we failed to find the concave impact of age on health Real Change. So we substitute age and age-squared to log (age). [2] We included seven occupations as mentioned to begin with, and only the coefficient of farmers is significant. So we include farmers only and rerun the regressions.

Note that our model is similar, but not the same as the traditional first-differences model. In Table 5-5, we can see that generally the impact of SES on real health change works in the same way as for SRH with fixed-effects, except for marriage which fails to be significant. Education has a significant positive impact on the change of health status, which means that individuals with more education are less likely to experience a negative change in their health. Self-reported health declines with age, and the speed of decline also increases with age. Household income is another important factor on determining individuals' health change as more household income decreases the possibility of negative changes. In addition, within the occupations, only being a farmer is closely linked to the negative change of health.

5.5.4. The Documented Healthy BMI Range – An Alternative Measure of Health

Because health has many dimensions, it is of interest to obtain results by using several health measures, although it is restricted by the data availability. In this section, we take the Healthy BMI calculated in equation (5.2) as the dependant variable, and run the regression with independent variables specified in equation (5.1). The Healthy BMI as shown in equation (5.2) measures the distance of the individuals' BMI values from the healthy range (19-25), and the greater the distance from either side the less healthy people are. The total value range of BMI is from 9 to 47, and Table 5-2 shows that the BMI is within the healthy range for 72% of the individuals (73% of the men and 70% of the women).

In addition to the regression with the Healthy BMI (full sample), the variables 'Healthy & Overweight' and 'Healthy & Underweight' are also separately included. Based on the specification of the dependant variable, a Tobit model is selected for its ability to account for the effects of censoring at the lower bound of the BMI risk ladder (i.e., at value 0). We further allow for random-effects as the estimation of a fixed-effects Tobit model is problematic (see e.g. Greene, 2004). Table 5-6 presents the estimation results.

The first two columns of Table 5-6 show coefficient estimates of the pooled and random-effects Tobit with the dependant variable of Healthy BMI. The likelihood ratio test shows that the panel-level variance component is important, thus we will focus on the results of random effects. Individuals' gender is a significant predictor of healthy BMIs, and men are less frequently overweight/underweight than women[94]. People are more likely to have a healthy BMI when they are getting older, but this is only until age 20. In other words, for the majority of our observations, older people are increasingly less likely to be in the healthy BMI range. As to the impact of SES, people with healthy BMIs are most common among individuals with lower household income. Overweight/underweight is overrepresented among administrators, whereas healthy BMIs are common among farmers. In addition, people's education attainment fails to have significant effect on their BMIs for the full sample.

The third and fourth columns of Table 5-6 show coefficient estimates of the pooled and random-effects Tobit with the selected samples on Healthy & Overweight. Not

[94] This is predicted on the assumption that the same healthy range is applicable for both men and women.

surprising, compared to the second column of Table 5-6, results in the fourth column show a clearer trend of the impact of age and SES on healthy BMIs. People are more likely to be overweight when they are getting older till they are approximately 60. Being overweight is a characteristic most common among married people, and together with the results in the last two columns of Table 5-6, married people are generally less likely to be underweight. People living in urban areas tend to be overweight, and this is consistent with the results of its impact on SRH. In addition, people's education plays a role in determining their BMIs as more educated people are less likely to be overweight. Compared to the results with the full sample Healthy BMI, the positive association between household income and being administrators and overweight is more robust. This is consistent with many other studies for developing countries. Reddy (1998) argued that this positive association is qualitatively different from the negative association characterizing contemporary Western populations. In the industrialized West, the richer upper strata of the populations, at least, are known to eat a more balanced diet, more possibly exercise during their leisure time and are more conscious of the need to check their weight. This to a large extent has been just the opposite in developing countries, such as China.

Table 5-6 The Impact of SES on Healthy BMI: (Random-Effects) Tobit Model

Explanatory Var.	**Full Sample**		**Healthy&Overweight**		**Healthy&Underweight**	
	Tobit	RE-Tobit	Tobit	RE-Tobit	Tobit	RE-Tobit
Gender	-0.431***	-0.343***	-0.953***	-0.677***	-0.069**	-0.046*
	(-8.93)	(-7.56)	(-11.10)	(-10.55)	(-1.97)	(-1.91)
Educ	-0.016**	-0.010	-0.032**	-0.019*	-0.002	-0.004
	(-1.99)	(-1.42)	(-2.39)	(-1.89)	(-0.40)	(-0.96)
Age	-0.041***	-0.020**	0.197***	0.164***	-0.126***	-0.090***
	(-3.55)	(-1.97)	(8.78)	(10.36)	(-15.60)	(-16.64)
Agesq	0.0007***	0.0005***	-0.0016***	-0.0013***	0.0015***	0.0011***
	(5.95)	(4.37)	(-6.40)	(-7.74)	(16.69)	(17.54)
Married	-0.110	-0.063	0.459***	0.299***	-0.248***	-0.139***
	(-1.49)	(-1.02)	(3.25)	(3.07)	(-4.92)	(-4.12)
Urban	0.079	0.091*	0.234**	0.205***	-0.050	-0.035
	(1.47)	(1.80)	(2.55)	(3.01)	(-1.22)	(-1.25)
Ladhinc	0.091***	0.068***	0.244***	0.165***	-0.060***	-0.039***
	(2.98)	(2.92)	(4.45)	(4.52)	(-2.72)	(-2.77)
Profes	0.296**	0.151	0.334*	0.200	0.113	0.062
	(2.55)	(1.56)	(1.78)	(1.51)	(1.17)	(0.96)
Adminis	0.497***	0.321***	0.706***	0.441***	-0.148	-0.078
	(4.09)	(3.26)	(3.70)	(3.34)	(-1.31)	(-1.07)
Staff	0.023	0.001	-0.078	-0.049	0.109	0.070
	(0.17)	(0.01)	(-0.36)	(-0.34)	(1.00)	(0.99)
Farmers	-0.609***	-0.415***	-1.656***	-1.066***	0.113**	0.077**
	(-8.49)	(-7.17)	(-13.18)	(-12.35)	(2.10)	(2.20)
Skilled	-0.028	-0.015	-0.039	0.01	0.010	0.002
	(-0.28)	(-0.20)	(-0.23)	(0.09)	(0.13)	(0.05)
Drivers	0.26	-0.019	0.663**	0.28	-0.063	-0.109
	(1.27)	(-0.12)	(1.99)	(1.24)	(-0.37)	(-0.99)
Service	0.180*	0.084	0.337**	0.183	-0.014	-0.025
	(1.73)	(1.04)	(1.98)	(1.59)	-(0.17)	(-0.48)
Reg. Dummies	yes	yes	yes	yes	yes	yes
Year Dummies	yes	yes	yes	yes	yes	yes
Log likelihood	-26970.73	-25240.07	-17027.71	-16236.79	-12117.40	-11754.57
Likelihood-ratio		3461.33		1580.83		725.68
No. of the Obs.	26366	26366	22931	22931	22336	22336

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions estimated by Tobit and random-effects Tobit (RE-Tobit).

5.5.5. An Alternative Measure of BMI and its Impact on SRH

As argued before, the BMI is more closely related to health behavior than SRH, and some researchers have also taken it as an important element of a life style (see, e.g. Contoyannis and Jones, 2004). The BMI is generally calculated as weight in kilograms over height in meters squared (weight(kg)/height(m)²), and it is currently the most commonly used method in the empirical work. But this formula appears to be somewhat arbitrary. Why divide weight by exactly the square of height? Why not the cube, why not simply the ratio itself? In order to investigate this more closely, it is necessary to consider how to define BMI. Considering for cross-sectional data for simplicity, we started by estimating the following equation:

$$\text{Log}\,(W_i) = \beta_0 + \beta_1 \log\,(H_i) + \boldsymbol{\beta_2 X_i} \qquad (5.12)$$

Where W_i and H_i denote the weight and height of the i'th individual, $\boldsymbol{X_i}$ is a vector of socio-economic characteristics and the log is to the base e. If we exponentiate both sides of equation (5.12), we get:

$$W_i = e^{\beta_0} H_i^{\beta_1} e^{\boldsymbol{\beta_2 X_i}} \qquad (5.13)$$

According to the literature, it suggests that an individual's Body Mass can be formulated by:

$$BMI_0 = W_i / H_i^{\beta_1} \tag{5.14}$$

where the subscript denotes that the equation (5.14) is linked to the 'standard way' of calculating BMI, indeed if β_1=2, then it is exactly the equation documented in the literature. A potential problem revolves around this measure in whether there are systematic differences in people's weight given their height. Thus whether an individual's BMI should be adjusted for gender, age and SES? Arguably it should if it is to reflect an unambiguous measure of BMI to be used in evaluating health. In this case, equation (5.14) should be amended to:

$$BMI_a = W_i / (H_i^{\beta_1} e^{\boldsymbol{\beta_2 X_i}}) \tag{5.15a}$$

Or viewed alternatively

$$e^{\boldsymbol{\beta_2 X_i}} BMI_0 = W_i / H_i^{\beta_1} \tag{5.15b}$$

Equation (5.15a) shows an amended BMI (BMI_a) based on gender, age and SES, and alternatively equation (5.15b) shows a 'standard' BMI which should be adjusted by gender, age and SES when interpreting it. $\mathbf{X}_i$ presents the gender, age and SES of *i*'th individual as in equation (5.12). In what follows we shall first base our estimates on equation (5.14) without the gender, age and SES variables as is consistent with the traditional approach. The estimated regression based on panel data with random-effects is:

$$\text{Log}(W_i) = 3.04 + 2.09\text{Log}(H_i)$$

$$(289.816) \quad (94.878)$$

$$R^2 = 0.40, n=26366 \tag{5.16}$$

Where the figures in parentheses represent *t*-statistics and the time subscript is implicit. From this we can calculate the BMI as:

$$BMI_0 = W_i / H_i^{2.09} \tag{5.17}$$

The coefficient is close to, but slightly higher than, the standard measure of 2.00 in the literature. Still, it provides some justification for it.

We now turn to consider the equations with gender, age and SES included, and we start by the inclusion of gender. We take $\text{Log}(W_i)$ as a dependant variable and run the regression on $\text{Log}(H_i)$ and gender, and we obtain a value for β_1 of 2.27. The whole regression results are as follows:

$$\text{Log}(W_i) = 2.97 - 0.029\text{gender} + 2.27\text{Log}(H_i)$$

$$(233.667) \quad (9.583) \quad (78.912)$$

$$R^2 = 0.40, n=26366 \tag{5.18}$$

Based on these results, we can calculate BMI by:

$$\mathrm{BMI}_1 = \mathrm{W}_i / (\mathrm{H}_i^{2.27}\exp(-0.029*\mathrm{gender})) \tag{5.19a}$$

Or

$$\exp(-0.029*\mathrm{gender})\mathrm{BMI}_0 = \mathrm{W}_i / \mathrm{H}_i^{2.27} \tag{5.19b}$$

Equations (5.19a) and (5.19b) could arguably contain a constant term, but β_0 reflects other influences and hence should not perform this role. But if we assume that people on average are healthy this could be used to normalize (5.19a). Equation (5.19a) shows that the coefficient on height has increased. To put it differently, the coefficient on height changes when account is taken of gender. We also see that for a given height and weight, the healthy BMI is different for men than women. Note that gender and age are both reflecting people's physical characteristics, and they should be both adjusted for, if we want to obtain an unambiguous measure of BMI to be used in evaluating health. Thus we now report a regression with both gender and age included, and we obtain a value for β_1 of 2.40. In this way we define the BMI as:

$$\mathrm{BMI}_2 = \mathrm{W}_i / (\mathrm{H}_i^{2.40}\exp(-0.039*\mathrm{gender} + 0.011*\mathrm{age} - 0.0001*\mathrm{age}^2)) \tag{5.20a}$$

Or

$$\exp(-0.039*\mathrm{gender} + 0.011*\mathrm{age} - 0.0001*\mathrm{age}^2)\mathrm{BMI}_0 = \mathrm{W}_i / \mathrm{H}_i^{2.40} \tag{5.20b}$$

We can divide the part (exp(.)) in equation (5.20a) into two components: exp(–0.039*gender) and exp(0.011*age – 0.0001*age^2), and we will look at them separately. From the coefficient of gender in equation (5.20a), it suggests that for a given height, weight and age, a BMI_2 value for a man (gender=1) is some 4% (= 1 – exp(-0.039)) lower than for a woman (gender=0). As to age, the coefficients indicate that the maximum value for exp($\beta_2$age + $\beta_3$age^2) is when age equals 55[95]. This suggests that BMI_2 values are greatest for people aged 55, i.e., given an individual's weight, height and gender, a BMI_2 value for someone who is 20, 40 or 70 is systematically lower than for someone who is 55. To sum up, for example, the BMI_2 value for a woman aged 50 is some 10% higher than a man aged 30[96]. This is consistent with the discussions in chapter 3 that the BMI values are age and gender dependent when used as an indicator of body fatness, and specifically they gradually increase with age, until after age 50 to 60 (see e.g., de Vasconcellos, 1994).

Finally we run the regression on Log(H_i) and all SES variables and from this way we obtain an estimate of β_1 of 2.24 as shown in Table 5-7. In this case, more interesting than the significance of height, gender and age, is the significance of the other variables. Education, household income and administrative and service occupations are all significant at the 1% level and urban is significant at the 5% level. In all cases weight increases with the variable, i.e. it increases with education and income and is higher for executives or those in the service sector and/or those living in towns. Hence to answer

---

95 $(e^{(0.011age-0.0001age^2)})' = e^{(0.011age-0.0001age^2)} * (0.011 - 0.0002age) = 0$, within this equation, $e^{(0.011age-0.0001age^2)}$ can not equal to 0, thus we can only let (0.011 – 0.0002*age*) = 0, and get age=55.

96 $(e^{(0.011*50-0.0001*50*50)} - e^{(-0.039+0.011*30-0.0001*30*30)})/e^{(-0.039+0.011*30-0.0001*30*30)}$ = 0.104 or 10.4%

the question posed earlier there are systematic differences in people's weight given their height which are unrelated to physical characteristics such as gender and age. We cannot say that the healthy values of BMI for more educated people, for example, should be lower or higher than less educated people, it depends upon their other physical characteristics. In addition, if height is correlated with education, urban, income, etc, then the exclusion of these variables in the equation will bias the coefficient on height[97]. In addition, it seems reasonable to suggest that all of these variables are ones which tend to be associated with less exercise. For example, people who live in towns will tend to have access to better public transport and walk less than people in rural areas. Similarly less well-educated people probably tend to have more manual jobs.

What does this imply about the measure of BMI? In particular should it be adjusted for people's socio-economic status as well as their physical characteristics – gender and age? The answer is no. There seems little reason why someone of a given age, gender, height and weight should have his optimal BMI adjusted because they are an executive or live in the city[98]. However, what it does suggest is that such socio-economic characteristics impact on people's BMI values and hence on health, and also that failing to take cognizance of socio-economic characteristics when estimating the relationship between weight and height can lead to biased results.

---

[97] Such correlations may exist because of inter-generational advantages, for example, the children of richer families tend to eat better during childhood and hence arguably become taller.
[98] Of course certain life styles are unhealthy which are linked to unhealthy BMI values. An executive life style, for instance, may be a stressful one – as indeed may be a coal miner's. That is not the issue. Instead we are asking whether the characteristics of a 12 stone, 6 foot, 35 year old man should have different interpretations vis-à-vis BMI values, depending on the individuals level of education or geographic location and on this we believe the answer is no.

Table 5-7 Determinants of BMI Values

Explanatory Variable	Log(weight)			
	RE-REG	RE-REG	RE-REG	RE-REG
Gender		-0.029***	-0.039***	-0.036***
		(-9.583)	(-13.252)	(-12.392)
Age			0.011***	0.011***
			(35.446)	(30.646)
Agesq			-0.0001***	-0.0001***
			(-27.440)	(-23.485)
Education				0.003***
				(10.646)
Married				-0.0004
				(-0.186)
Urban				0.005**
				(2.223)
Ladhinc				0.002***
				(2.766)
Profes				0.0001
				(0.024)
Admins				0.011***
				(3.627)
Staff				0.002
				(0.708)
Farmer				-0.016***
				(-8.739)
Skilled				0.0004
				(0.193)
Driver				0.009*
				(1.955)
Service				0.009***
				(3.658)
Log(height)	2.094***	2.272***	2.395***	2.247***
	(94.878)	(78.912)	(83.827)	(76.840)
R-squared	0.40	0.40	0.44	0.45
No. of the Obs.	26366	26366	26366	26366

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions are estimated based on panel data with random-effects (RE-REG).

Recalculating the Healthy Range

So far we have developed a little further the issue on how to define BMI. We now face the question of examining the 'healthy range' based on the new measures of BMI and estimating its impact on SRH together with SES. We will start with BMI_0 derived from equation (5.17), which is close to that used in the literature. However, our main focus is on the relatively complex one (BMI_2) taking into account gender and age from equation (5.20a). The total value ranges of BMI measure based on equation (5.17) and (5.20a) is (8-42) and (5-33), but the main observations are within (15-35) and (10-25) respectively. For BMI_0, we construct two variables BMI_0^1 and BMI_0^2 as follows:

$$BMI_0^1 = (\alpha_L - BMI_0), \text{ operative if } BMI_0 < \alpha_L$$

$$BMI_0^2 = (BMI_0 - \alpha_H), \text{ operative if } BMI_0 > \alpha_H \qquad (5.21)$$

and similarly for BMI_2:

$$BMI_2^1 = (\alpha_L - BMI_2), \text{ operative if } BMI_2 < \alpha_L$$

$$BMI_2^2 = (BMI_2 - \alpha_H), \text{ operative if } BMI_2 > \alpha_H \qquad (5.22)$$

The range α_L - α_H is what we term the 'healthy range'. People outside this range are unhealthy, and the greater the distance the less healthy they are. We identify the critical points α_L and α_H through regression analysis. Self-reported health is regressed on (BMI_0^1, BMI_0^2) and (BMI_2^1, BMI_2^2) by an iterative search technique where the critical values for α_L and α_H included all possible combinations from (15-35) and (10-25), which amounted to 231 and 136 regressions based on equation (5.17) and (5.20a) respectively. We choose the optimal combination on the basis of i) significantly negative coefficients for both (BMI_0^1, BMI_0^2) and (BMI_2^1, BMI_2^2), and ii) by identifying the best fit by the highest log likelihood ratio. Based on this, two BMI healthy ranges have been identified: using BMI_0 as in equation (5.17), we find critical values for α_L of 22 and α_H of 27[99]; Using the modified BMI (BMI_2) in equation (5.20a), we identify critical values of 15 to 19. Table 5-8 shows the random-effects ordered probit results[100] which underlie these calculations.

At this stage, the 'best' results based on the log likelihood appear marginally to be those based on the standard measure. Note that the healthy BMI range based on BMI_0 in equation (5.17), i.e., (22-27) per se is higher than (19-25) as documented in literature. The potential reason perhaps is that we are using different measure of health. People with BMIs above 25 are considered overweight or 'at risk', however, these risks are mainly relevant to obesity-related diseases, instead of self-report measure of health which we use in our estimate. In addition, there is some suggestion that BMI_0^2 and BMI_2^2 may have greater impact on health than BMI_0^1 and BMI_2^1, i.e. being underweight has less severe consequences than being overweight. Although this is not something we develop further in this section. The significance of the other variables is largely as before.

---

[99] For a comparison purpose, we also use the traditional measure of BMI based on the formula weight(kg)/height(m)², we find critical values for α_L of 23 and α_H of 28.

[100] Note that we obtain the same healthy ranges based on pooled and random-effects results.

Table 5-8 The Optimal Calculations for (BMI_0^1, BMI_0^2) and (BMI_2^1, BMI_2^2)

Explanatory Variable	Self-reported Health	
	RE-OP1	RE-OP2
Gender	0.253***	0.224***
	(7.946)	(7.067)
Education	0.024***	0.024***
	(4.737)	(4.741)
Age	-0.043***	-0.025***
	(-5.761)	(-3.363)
Agesq	0.00003	-0.0001*
	(0.339)	(-1.677)
Married	-0.011	-0.004
	(-0.244)	(-0.087)
Urban	-0.385***	-0.385***
	(-10.578)	(-10.582)
Ladhinc	0.109***	0.108***
	(5.945)	(5.925)
Profes	-0.123*	-0.118
	(-1.648)	(-1.581)
Admins	0.186**	0.190**
	(2.398)	(2.449)
Staff	0.148*	0.156*
	(1.792)	(1.884)
Farmer	-0.091**	-0.094**
	(-2.019)	(-2.070)
Skilled	0.171***	0.177***
	(2.725)	(2.811)
Driver	0.292**	0.304**
	(2.128)	(2.223)
Service	0.108	0.108
	(1.582)	(1.574)
BMI_0^1/ BMI_2^1	-0.123***	-0.192***
	(-13.453)	(-13.165)
BMI_0^2 / BMI_2^2	-0.103***	-0.119***
	(-4.441)	(-3.476)
Regional dummies	Yes	Yes
Year dummies	Yes	Yes
Log likelihood	-25497.53	-25504.45
Restricted log likelihood†	-25658.92	-25666.20
No. of the Obs.	26366	26366

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. † Log-likelihood ratio is calculated by $2(L_{unrestricted} – L_{restricted})$. Regressions are estimated by random-effects ordered probit. BMI_0^1 and BMI_0^2 are defined by equation (5.21), and BMI_2^1 and BMI_2^2 are defined by equation (5.22). The results of first column (RE-OP1) are based on (BMI_0^1, BMI_0^2), i.e. (22-27). The results of second column (RE-OP2) are based on (BMI_2^1, BMI_2^2), i.e. (15-19).

Moderating Influences on an Adverse BMI

We now construct a New BMI variable based on BMI_2 with the optimal value range of 15 (α_L) and 19 (α_H) obtained from equation (5.22). It is defined as the sum of BMI_2^1 and BMI_2^2. The distance between the BMI_2 values and either 15 (α_L) or 19 (α_H) represents a deviation of BMI_2 from its healthy range. We attempt to estimate how this healthy range is associated with SES in its impact on health. We construct the following interaction terms: New BMI*gender, New BMI*age, New BMI*age^2, New BMI*education, and New BMI*log(household income). Being as we are using panel data, we also include year dummies. Table 5-9 shows the results.

Table 5-9 The Impact of New BMI together with SES on SRH

Explanatory Var.	**Self-reported Health**			
	OProbit	RE-OProb	OProbit	RE-OProb
Gender	0.205^{***}	0.223^{***}	0.261^{***}	0.286^{***}
	(7.823)	(7.053)	(7.665)	(7.135)
Education	0.021^{***}	0.024^{***}	0.027^{***}	0.031^{***}
	(4.789)	(4.708)	(5.223)	(5.084)
Age	-0.024^{***}	-0.026^{***}	-0.013^{*}	-0.012
	(-3.811)	(-3.440)	(-1.744)	(-1.368)
Agesq	-0.0001	-0.0001	-0.0002^{**}	-0.0003^{***}
	(-1.512)	(-1.623)	(-2.513)	(-2.752)
Married	0.012	-0.003	0.012	-0.002
	(0.290)	(-0.060)	(0.306)	(-0.042)
Urban	-0.352^{***}	-0.384^{***}	-0.354^{***}	-0.386^{***}
	(-11.818)	(-10.557)	(-11.869)	(-10.604)
Ladhinc	0.109^{***}	0.109^{***}	0.084^{***}	0.084^{***}
	(6.627)	(5.982)	(3.974)	(3.603)
Profes	-0.119^{*}	-0.119	-0.127^{*}	-0.126^{*}
	(-1.781)	(-1.599)	(-1.899)	(-1.696)

Table 5-9 Continued

Explanatory Var.	Self-reported Health			
	OProbit	RE-OProb	OProbit	RE-OProb
Admins	0.178**	0.191**	0.166**	0.178**
	(2.536)	(2.462)	(2.354)	(2.291)
Staff	0.143*	0.154*	0.139*	0.149*
	(1.885)	(1.868)	(1.822)	(1.799)
Farmer	-0.084**	-0.098**	-0.083**	-0.096**
	(-2.137)	(-2.168)	(-2.101)	(-2.119)
Skilled	0.158***	0.176***	0.156***	0.175***
	(2.770)	(2.793)	(2.750)	(2.784)
Driver	0.294***	0.304**	0.285**	0.294**
	(2.582)	(2.220)	(2.501)	(2.147)
Service	0.116**	0.108	0.117**	0.109
	(1.984)	(1.579)	(2.009)	(1.600)
BMINEW	-0.160***	-0.182***	-0.029	0.024
	(-13.101)	(-13.264)	(-0.194)	(0.143)
BMI*gender			-0.062**	-0.070**
			(-2.531)	(-2.535)
BMI*age			-0.013***	-0.016***
			(-2.765)	(-3.071)
BMI*agesq			0.0001**	0.0002***
			(2.495)	(2.820)
BMI*education			-0.008**	-0.008**
			(-2.263)	(-2.106)
BMI*Ladhinc			0.028**	0.028*
			(1.972)	(1.790)
Regional dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Log likelihood	-25668.41	-25506.37	-25655.74	-25493.98
Restricted log likelihood	-27238.45	-25668.41	-27238.45	-25655.74
No. of the Obs.	26366	26366	26366	26366

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions are estimated by ordered probit (OProbit) and random-effects ordered probit (RE-OProb).

The first two columns of Table 5-9 show the estimated coefficients for the ordered probit models based on pooled and random-effects specifications with the inclusion of the New BMI variable. The coefficient of New BMI is significantly negative as expected, which shows that when a person's BMI (based on the equation of BMI_2) value moves away from the healthy range, he/she tends to be less healthy. We now look at whether the impact of this New BMI on health changes by different SES. Through the last two columns of Table 5-9, we can see that the impact of New BMI on health is affected by gender, age, education and income[101]. Males are basically healthier than females, but this is changed when account is taken of the interaction with the New BMI. The coefficient of New BMI*gender is significantly negative, which indicates that, for a given BMI, there is a greater adverse impact on health for a man than a woman. Similarly, the impact of New BMI on health also depends on age, and more specifically, it is convex, first increasing and then decreasing after reaching a certain age. Based on the random-effects results in the last column of Table 5-9, the impact of New BMI on health is highest for someone aged 40. In addition, higher income neutralizes to some extent an adverse New BMI. Finally and interestingly, a given value of New BMI has a greater adverse impact on health with more education. Viewed in another light, we have concluded that educated people tend to be healthier than less educated people, presumably because they have increased awareness of what constitutes 'healthy living'. An adverse BMI with more education on health, however, suggests that they are not utilizing this knowledge, thus reducing the 'educational advantage' in health. These results give further credence to the view that in interpreting BMI we need to make distinctions in the basis of age and gender. In addition, there is the possibility that the consequences of an adverse BMI can be reduced by greater income, possibly because of

[101] Note that we have also included the New BMI interactive variables with married, urban and occupations, however, all of them fail to be significant. Results do not report in Table 5-9.

better access to medical facilities.

5.6. Conclusions

It has been documented extensively that the impact of SES on health exists, and the most heavily researched topics focus on the relationships between education, income and health. Few studies apply datasets from outside the US/Europe or allow for individual specific heterogeneity over time, which we do in this chapter. We used a four-year (1991, 1993, 1997 and 2000) panel data set to estimate the impact of SES on self-reported health and BMI, as an alternative measure of health. We apply recently-developed statistical methods for estimating an ordered logit model with fixed-effects as well as an ordered probit model with random-effects. As to the impact of education and income on health, one of the main concerns of this chapter, entering individual fixed-effects do not make a significant difference. Specifically the income, the magnitude and significance of the coefficients in random-effects and fixed-effects models are very close. Thus the random-effects model has not been far away from the mark in the respect of unobservable heterogeneity. In addition, as with many other countries, we have also confirmed the impact of gender, age, living areas, etc. on health. To further deal with potential simultaneity between education and health, we used an approach based on the relationship between SES and health changes, which is similar, but not the same as the traditional first-differences model. Based on this, we find the SES pattern for people whose health statuses are getting worse. Further as a robustness check, this impact of SES on health change works in the same way as for SRH with fixed-effects, except for marriage which fails to be significant.

BMI is closely related to nutrition and health behavior as documented in the literature. In this respect, the estimation of the relationship between SES and BMI is even more important. We find consistent results with many other studies for developing countries on the positive association between income and certain occupations, such as administrators and farmers, and being overweight. As to the impact of education, we find that being overweight is least common among more highly educated. Thus education impacts positively not only on health and health change, but also on the possibility of individuals' being in the documented healthy BMI range. Furthermore, we have confirmed that the concept of the BMI is a valid one in the literature, and have further argued that it is reasonable that it should be calculated in such a manner as adjusts for gender and age. We examine the healthy range of this adjusted measure of BMI based on the self-reported health. Note that we are using a different measure of BMI, thus the comparison of the healthy range with other studies is rather difficult. We admit the possibility that the self-reported measure may not be as accurate as the morbidity and mortality data which is mostly applied in this context, still, it is of interest in its own right. Importantly, variables such as income can modify the impact of an adverse BMI, possibly because of better access to medical facilities.

Chapter 6 SES, Social Capital and Subjective Well-being

6.1. Introduction

The study of subjective well-being[102] (SWB) has been an area of increasing interest over the last three decades by psychologists, sociologists and economists. Psychologists focus upon factors that keep one from being depressed and factors that lead one to becoming elated (see, e.g., Eddington and Shuman, 2005). Economists favor the concept of utility to explain individuals' choices and behavior. Together with sociologists, they are mostly interested in establishing the determinants of SWB. Recent reviews can be found in Frey and Stutzer (2002) and Layard (2005). This is an important area of research, according to Hudson (2006), which economists have been somewhat ambivalent about. One of the hypotheses which has emerged from the recent literature is the impact of social capital on well-being. Specifically at the level of individual, associational involvement and interpersonal trust have been found to be related to self-assessments of individual well-being (see, e.g., Putman, 2000).

Social capital has become a popular topic in the past decade and research linking it with

[102] Subjective well-being is the psychological term for "happiness", and within the literature, the terms are used interchangeably (Eddington and Shuman, 2005).

subjective well-being has come fast and furious. The most robust correlate of social capital variables across individuals is years of education completed, and more broadly, education closely linked with both social capital and SWB. Most of the relevant work has been conducted in the western countries, particular in the United Stated, and comparatively fewer studies have been done in other contexts. Because culture plays a potentially important role in this area, the relatively few studies based on Chinese participants is an obvious limitation when trying to reach conclusions with respect to China. With the deepening of Chinese economic reform and the possible increase in the quality of life, this topic will almost certainly receive more attention. Thus it is important to fill this gap and conduct a study based on Chinese culture to shed light on cross-cultural differences. This will be, as far as we are aware, the first time the impact of social capital (interpersonal trust and participation in voluntary associations in particular) has on SWB has been analyzed in the Chinese context.

We aim to examine whether, and to what extent, SES and social capital have impacts on happiness and life satisfaction for China, as they do in other countries, using cross-sectional dataset of World Values Surveys (1990, 1995-1997, 1999-2000). It will be followed by an analysis of the determinants of social capital. As throughout the thesis, the main area of concern lies with the impact of education on both social capital and SWB. We will also look at the difference between happiness and life satisfaction.

This chapter is organised as follows: In the next section, we provide a comprehensive literature review on both SWB and social capital. In section 3, we describe the theoretical foundation and data specification. Section 4 presents the empirical results. This will be in two stages. Firstly the determinants, mainly SES and social capital, on SWB and secondly the determinants of social capital. Finally we conclude this chapter.

6.2. Literatures Review

6.2.1. Happiness

Scientific definitions of SWB recognize an affective and a cognitive component of well-being. As Frey and Stutzer (2002) explained that "*SWB is an attitude consisting of two basic aspects of cognition and affect. 'Affect' is the label attached to moods and emotions. Affect represents people's instant evaluation of the events that occur in their lives. The cognitive component refers to the rational or intellectual aspects of SWB and it is usually assessed with measures of satisfaction*". Thus people are said to have high level of SWB if they are satisfied with their life-conditions, and experience frequent positive emotions and infrequent negative emotions (Eddington and Shuman, 2005).

Several dimensions in current research have been applied to SWB including overall feelings of happiness, life satisfaction and some domains of life satisfaction, such as job satisfaction, family satisfaction, and financial satisfaction, etc. Many studies, either from economics, sociology or psychology, have shown that demographic factors, such as age and gender, are important determinants of happiness and life satisfaction. Arygle (1987) mentioned that if the emphasis is on satisfaction or cognitive appraisal of well-being, there is a definite increase with age. Recently a U-shaped relationship has been found between age and happiness, with the young and old more satisfied than the middle aged (Namazie and Sanfey, 1999; Donovan and Halpern, 2002; Hayo and Seifert, 2003; Hudson, 2006). Donovan and Halpern (2002) argued that there are significant national differences. They mentioned that life satisfaction is highest amongst the over 65s in Japan, and by contrast, satisfaction is lower amongst older generations in Hungary[103]. In some western countries, such as Denmark and Britain, the difference between age groups is modest throughout the age range. Gender is a factor typically included in this context, and the levels of happiness and life satisfaction are generally found to be higher among women than men[104] (Donovan and Halpern, 2002; Blanchflower and Oswald, 2004). However, Frey and Stutzer (2002) argued that this difference is not large and has tended to disappear in recent years.

Apart from these, it is important to distinguish the following types of determinants.

---

[103] This may be linked with the transition process from planned to market economy.

[104] Donovan and Halpern (2002) further argued that the notable exception is Russia, where women are less happy than men. Also, in South American countries men and women appear to be equally happy.

Economic Factors - Income and Unemployment

Economic factors have been shown to be important contributors to individuals' subjective well-being (Clark and Oswald, 1994; Blanchflower and Oswald, 2004). In particular economists have been interested in establishing the relationships between income, unemployment and life satisfaction. The basic starting point for many analyses is income. The relationship between income and satisfaction is less clear, and the most widely accepted viewpoint is that income does matter, but not very much (Easterlin, 1995; Oswald, 1997; Frey and Stutzer, 2002). The former indicates that people with higher incomes are more satisfied than those with lower incomes, and an increase in personal income does bring higher levels of satisfaction. For example, Frijters, *et al.* (2004) have concluded that around 35-40% of the increase in life satisfaction in East Germany was attributable to the large increase in real household incomes. Further, investigations of longitudinal data on the effects of windfalls (lottery wins and inheritances) showed that the income does have a causal effect upon life satisfaction (Gardner and Oswald, 2001). The latter has mainly led to interest in the role of relative rather than absolute income in determining life satisfaction (see e.g. Clark and Oswald, 1996; McBride, 2001). The relative income discussion has been confirmed by Ferrer-i-Carbonell (2005), who concluded that the larger an individual's own income in comparison with the income of the reference group, the happier the individual will be. Stutzer (2004) approached this in a different way, i.e. the higher the ratio of aspired income to actual income, the less satisfied people are with their lives.

Some researchers have reported on the effects of social class, which is closely linked to relative income, but is mainly based on different positions at work. Arygle (1987) pointed out that to a person in a certain occupation, satisfaction depends more on comparisons with what other people have, and on comparisons with their own past experience. In addition, there is a firm consensus based on both cross-sectional and longitudinal data that unemployment leads to a substantial loss of life satisfaction regardless of the exact definition of life satisfaction (Clark and Oswald, 1994; Frey and Stutzer, 2002; Frijters, *et al.*, 2004). Specifically, Frey and Stutzer (2002) concluded that the main direction of causality appears to run from unemployment to less satisfaction. This linkage between unemployment and life satisfaction is not just simply due to the loss of income. Frey and Stutzer argued that unemployment needs to be seen in a wider context, *the fate of being unemployed is less depressing when one compares oneself with others who are also out of work.* Thus the significance of being unemployed per se may be explained by a combination of the loss of self-esteem and the change of the norms of one's social group. There is another possibility, as argued by Hudson (2006), related to an aspirational standard of living. Regardless of their income whilst unemployed, it will have fallen compared to what it was when the individual was in work. In addition, given their income whilst unemployed they will be further from their aspirations than others on a similar income who are not unemployed.

Health and Choice or Control of One's Own Life

The existence of a significant linkage between health and SWB is well-documented. Helliwell (2003) reported that in the long run, a 1% increase in average self-reported health status is associated with just over 1% increase in SWB. However, it is important to note that strong correlation between health and SWB exists for self-reported health measures (Donovan and Halpern, 2002), but not for objective health ratings (see e.g., Diener, *et al.*, 1999). Therefore, perceptions of health appear to be more important than objective health in their effects on SWB although of course there is a linkage between these two. A possible explanation for this, argued by Eddington and Shuman (2005), is that people in poor health downplay the importance of their health when evaluating their life satisfaction, and also people use cognitive coping strategies that promote a positive image of their health condition. Apart from this, it has also been suggested that self-reported health partly reflects optimism or general sense of control (Bobak, *et al.*, 1998). Individuals whose personalities are inherently more optimistic are more likely to give positive assessments of their health status and their SWB. As Helliwell (2003) argued although a large fraction of long-term interpersonal differences in personality is genetic in nature, optimism does create real linkages with health outcomes. However, in his more recent paper, Helliwell (2006) argued that with the inclusion of a variable which captures a standard measure of personality, the coefficient on subjective health drops only slightly[105].

Control is important to psychological functioning (Skinner, 1996), and as suggested by Syme (1989), the concept of control may integrate many aspects of the social and psychosocial environment, including material, social and psychological well-being. There is a rich psychological literature showing that low perceived control, low ratings of self-efficiency and many other related concepts (see e.g., Skinner, 1996) are related to health status and behaviours, as the more general indicators of well-being (see e.g., Bobak, *et al.*, 1998). Apart from the psychological literature, Peterson and Stunkard (1989) argued that control plays a crucial role in people's choice of health behaviour. In social economics, the importance of a sense of control has been seen to have an impact on life satisfaction (Donovan and Halpern, 2002). This is not surprising. Loss of control constrains the welfare optimization problem, and should lead to sub-optimal outcomes.

6.2.2. Social Capital and Human Capital

A growing literature has focused on the impact of human capital and social capital on SWB. Studies which examine the linkage between human capital, specifically years of education completed, and SWB are mainly focused on the tendency for the more highly educated to have higher incomes, better health and more recently, more social capital (see

[105] From 0.54 to 0.49. For detailed discussions see Helliwell (2006).

e.g., Donovan and Halpern, 2002; Hayo and Seifert, 2003). In addition, Scitovsky (1976) argued that education allows people to take more advantage of activities which generate happiness or positive emotions, including music, painting and literature, etc., this effect would be a direct one independent of any affect on income or health. Like human capital, social capital provides important benefits to individuals and societies, and perhaps the most convincing evidence of positive impact of social capital lies in the area of personal health. Large numbers of cross-sectional studies have reported the strong association between the size and quality of people's social networks and their health, with people who are less socially isolated and more involved in social and civic activities tending to have better health. For example, Woolock (2001) argued that there is very strong evidence of powerful health effects of social connectedness, and the evidence is strong not only in the United States, but also in Finland, Japan and other countries. Controlling for a person's blood chemistry, age, gender, exercise or not, and for all other risk factors, his/her chance of dying over the course of the next year is cut in half by joining one group, and cut to a quarter by joining two groups. Putnam (2001) further argued that it is not that people who are healthy become joiners, it is from joining to health.

What are the direct effects of social capital on subjective well-being? Putnam (2000) confirmed marriage to be the most beneficial form of social interaction. Many researchers believed that marriage serves as a buffer against the hardships of life and it provides emotional support which produces positive states of well-being (Eddington and Shuman, 2005). The married are very substantially happier than the unmarried, and those living together are happier than those living alone (Donovan and Halpern, 2002). In addition, as some studies have indicated, this linkage appears to be different for men and women, and also at their different ages (Namazie and Sanfey, 1999; Donovan and Halpern, 2002; Dockery, 2004). Other measures of social connection also have significant effects on happiness, like being a member of voluntary associations (Argyle, 1987; Helliwell, 2003). Another factor which needs to be mentioned is dyadic trust, which is closely linked to inter-personal trust or social trust. Sometimes measures of trust are included as part of what is meant by social capital, while other researchers prefer to treat trust as something that is generated and supported by a more valuable sort of social capital (Helliwell, 2001; 2003). The preliminary conclusion is that people who trust others are generally happier and more satisfied with their lives (Donovan and Halpern, 2002). As Helliwell (2001) explained, high levels of inter-personal trust, particularly to the extent that this trust is matched by trustworthy behavior in others, do make many aspects of life more enjoyable and more productive, in part by reducing the costs of dealing with risk and uncertainty. In addition, Hudson (2006) specially focused on institutional trust and argued for an impact of institutions upon happiness as measured by their intermediary impact upon individual trust.

It is instructive to compare the relative importance of income on happiness with that of human and social capital. Putnam (2000) used a large sample from the DDB Needham

Life Style Survey to simultaneously assess the effects of education, income, and various kinds of social and civic engagement, so that each effect is an estimate of the marginal impact flowing directly rather than through one of the other. According to his results, four years of education has the happiness equivalent of a doubling of income, and further, a significant association between social engagement and happiness, with monthly club meetings and monthly volunteering both having the happiness equivalent of four extra years of schooling or a doubling of income. Evidence from the United States and Britain reviewed by Blanchflower and Oswald (2000) supported the view that social connectedness may be more important for happiness than education and income - at least for average to above-average levels of income. Myers (1999) found similar results for the United States on the impact of close personal relations and social connections on subjective reporting of quality of life.

Evidence has shown that social capital has an impact on SWB. However, as argued in chapter 3, these are not free of the risk of the confounding role of personality differences. For example, Ferrer-i-Carbonell and Frijters (2004) examined the robustness of the literature findings on the determinants of happiness. Using panel data, they argued that allowing for fixed personality traits does change results substantially. More specifically, optimists will tend to be more satisfied with their lives. They may well have 'rosy evaluations' of their health status. They may also be more likely to be married, and spend time with their neighbours, all of which have been found to be positively correlated with life satisfaction. However, according to Helliwell (2006), the 2003 Canadian General Social Survey (GSS) provides a good candidate measure, a 'mastery scale' based on answers to several questions designed to measure psychological coping resources. The results based on GSS show that the mastery variable is highly significant. There are no changes in the marital status coefficients when a personality-based variable is added for each individual; and the coefficients on other social capital variables, such as the trust placed in neighbours, are unaffected by the inclusion of the personality variable. These results indicated that these relationships are not solely or even predominantly attributable to personality differences.

What we have mentioned so far is the linkages between human and social capital and SWB. There is also a two-way relationship between social capital and human capital. There is a positive association even at a cross-country level, for example, between social and civic engagement and trust on the one hand, and levels of education on the other (OECD, 2001). As argued in chapter 3, the most robust correlate of social capital variables across individuals is years of education. According to Glaeser (2001), the raw correlation of years of education with membership in organizations is 34 percent in the GSS. Using the World Values Survey (an international version of the GSS), Glaeser, *et al*. (2000) found a positive relationship in almost every country. There is also an extremely strong connection between education and social trust. As Nie, *et al*. (1996) argued, individuals' trust levels are indeed pushed upward by both individual and average education levels.

6.2.3. Related Analysis on Happiness and Life Satisfaction for China

As mentioned before, relatively few studies have been done in the Chinese context or indeed in developing countries in general. Existing studies for China in this area are mostly focused on Hong Kong, one of the special administrative regions of China. These studies cover relatively broad issues on the quality of life, and with respect to specified age cohorts, such as adolescents and older adults (Shek, *et al.*, 2005; Cheung, *et al.*, 2005). One study by Shen and Choy (2005) focused on the comparison of the social well-being of Hong Kong residents before and after the change of sovereignty (on July 1997) by analyzing six sets of survey data between 1990 and 2001. They concluded that the residents were most satisfied in 1997, and the sovereignty change had major impacts on many life domains. Another study by Lu and Lin (1998) on Taiwan, China, examined four family roles: spouse role, parental role, filial role and worker role and their relationships with SWB, and reported that based on the traditional Chinese culture, role experiences were important to adult happiness.

Two recent studies have focused on mainland China. One study is by Zhang (2005), who looked at 1347 mainland Chinese across three generations aged from 14 to 88. Zhang estimated whether collective self-esteem had added value in prediction of life satisfaction beyond established predictors, i.e., demographic predictors, personality predictors and individual self-esteem. The results indicate that collective self-esteem added 1% explained variance to general life satisfaction and 3% explained variance to life domain satisfaction beyond those predictors. In a further study, Knight and Gunatilaka (2007) linked the literature on rural-urban migration and the literature on subjective well-being and examined the question of why rural-urban migrant households settled in urban China have an average happiness score lower than that of rural households. Based on a 2002 National Household Survey, their results indicated that rural-urban migrants' high aspirations in relation to achievement, influenced by reference groups, specifically make for unhappiness. One reason for the lack of studies on mainland China may partly be because of the lack of available datasets.

6.3. Model and Data Specification

6.3.1. Theoretical Foundation

As mentioned before, scientific definitions of SWB recognize an affective and a cognitive component of well-being. Thus it is not easy to formalize it. The following equation reflects a general approach often adopted in the economics literature (Frey and Stutzer, 2002; Blanchflower and Oswald, 2004):

$$W = H\,[U(\mathbf{X}, t)] + \varepsilon \qquad (6.1)$$

where W denotes the self-reported number or level of well-being (either on a cardinal scale or on an ordinal happiness scale, such as “very happy”, “quite happy”, etc.). U (…) is to be thought of as the person’s true well-being or utility, and is observable only by the individual themselves. Its structure is hidden from any other individuals. **X** denotes the determinants of reported subjective well-being. *t* indicates that the relationship between **X** and subjective well-being may vary over time, and ε is an error term which serves to capture other hidden factors influencing the difference between actual well-being and reported well-being. H [...] is a continuous non-differentiable function relating actual to reported subjective well-being, and the function H [...] rises as U increases. Blanchflower and Oswald (2004) concluded that the structure of equation (6.1) is suitable for estimation as an ordered probit. Thus, ‘true’ utility is the latent variable, and the subjectivity of responses can be thought of as being captured by the error term.

Ferrer-I-Carbonell and Frijters (2004) examined the robustness of the literature findings on the determinants of happiness. There are two main methodological issues: i) whether to treat reported happiness levels as cardinal (as psychologists generally do) or ordinal (as economists generally do); and ii) the influence of the unobserved determinants of happiness. As to the former issue, the authors argued that the results are not sensitive to the choice between OLS and latent variable methods. Blanchflower and Oswald (2004) also reported similar results between these two. With respect to the second issue, the results are sensitive to individual fixed personality traits. The implications are, for example, that having a personality which is conducive to happiness is also associated with having high income, being healthy, and being married. Although this is not consistent with Helliwell's (2006) argument in which these relationships are not solely or even predominantly attributable to personality differences, possibilities may still exist. Given that we lack panel information on happiness over time, the results thus are more illustrative than conclusive.

In our analysis, **X** will include SES, social capital variables and a set of demographic variables which are generally included in this sort of work. Specifically, we will give attention to social capital variables including participation in voluntary associations, party membership and social trust. These impact on SWB in two possible ways: indirectly those effects of social capital may flow through health, or even economic outcomes to well-being; and the direct linkage from social capital to well-being is related to affecting the efficiency of the individuals' welfare maximization problem, and people do gain significantly from social interactions in the social domain. Take social trust for example, failure to trust people means an inability to make use of potential help. A very simple example of this would be the reluctance to use a taxi because of a distrust of taxi drivers.

We will also be distinguishing between happiness and life satisfaction. Some researchers have argued for a similar structure of the happiness and life-satisfaction equations (see, e.g., Blanchflower and Oswald, 2004), as well as the possible difference which happiness is more responsive to the short-term circumstances compared to the life satisfaction (Helliwell, 2003). In terms of the earlier discussion we regard the former as reflecting 'affect' and the latter more cognitive or underlying level of happiness. In general, it seems reasonable to hypothesize a causal linkage from life satisfaction to happiness, although a reverse, probably weaker link from happiness to reported life satisfaction is possible. In addition, variables which facilitate an ability to 'enjoy life' may also impact on happiness.

6.3.2. The Data

We use the cross-sectional microdata of World Values Survey (1990, 1995-1997, 1999-2000), which is designed to enable a cross-national, cross-cultural comparison of values and norms on a wide variety of topics including economic conditions, social attitudes, social networks, quality of life, trust levels, and traditional values. We select the China part only, and estimate what we believe to be some of the first micro-econometric equations of happiness and life satisfaction for mainland China.

There are two dependant variables in our estimate to measure subjective well-being: how happy people feel and how satisfied with their lives. The former is based on the question of “Taking all things together, would you say you are......?” Possible responses are: i) very happy, ii) quite happy, iii) not very happy and iv) not at all happy[106]. A fifth possibility was ‘don’t know’, and the regression analysis excludes it. The latter is based on the question of “All things considered, how satisfied are you with your life as a whole?” The responses are composed of 10 scales with only two extreme values, where ‘1’ represents “dissatisfied” and ‘10’ represents “satisfied”. Not unexpectedly, 56.1% of survey respondents place themselves in the category of “quite happy (quoted ‘3’)”, and also a relatively high general life satisfaction with an average of ‘6.9’ out of 10 points. Figure 6-1 shows the histograms of happiness and life satisfaction, and Table 6-1 shows cross-tabulations between them. From the cross-tabulation between happiness and life satisfaction, we can see that within the category of people who choose ‘very happy’ (quoted ‘4’), the highest percentage (29.0%) of them tend to be the most satisfied with their lives (quoted ‘10’), whereas people who are in the category of the most satisfied with their lives (quoted ‘10’) tend to be just ‘quite happy’ (quoted ‘3’, 47.7%). Clearly these two concepts are closely correlated, but not identical. It is thus necessary to look at both measures to obtain a complete view of SWB for China.

Following the literature review and theoretical foundation, the independent variables in our regression analysis will include SES, social capital variables and a set of

[106] We have reversed the scale of the happiness measure in the regression to emphasize that higher numbers correspond to higher level of happiness.

demographic variables. The respondents were chosen by means of the random sampling selection covering 24 provinces[107], almost all the provinces of mainland China. The surveys use different codes[107] for regions between 1990, 1995-1997 and 1999-2000, thus it gives ambiguous information on the inclusion of provinces in each year. Hence, the regional dummies are not included in the regressions. We restrict our sample to those between 18 and 65. A large number of (around 81.8%) respondents are married and it is the included category for marital status. The other category combines the never married, those current widowed or separated, and those divorced. In addition, there is a slightly lower percentage (around 43.9%) of females than males in the survey. After also excluding observations with less than full information, it provides 2657 observations altogether. All the variables are defined in Table 6-2.

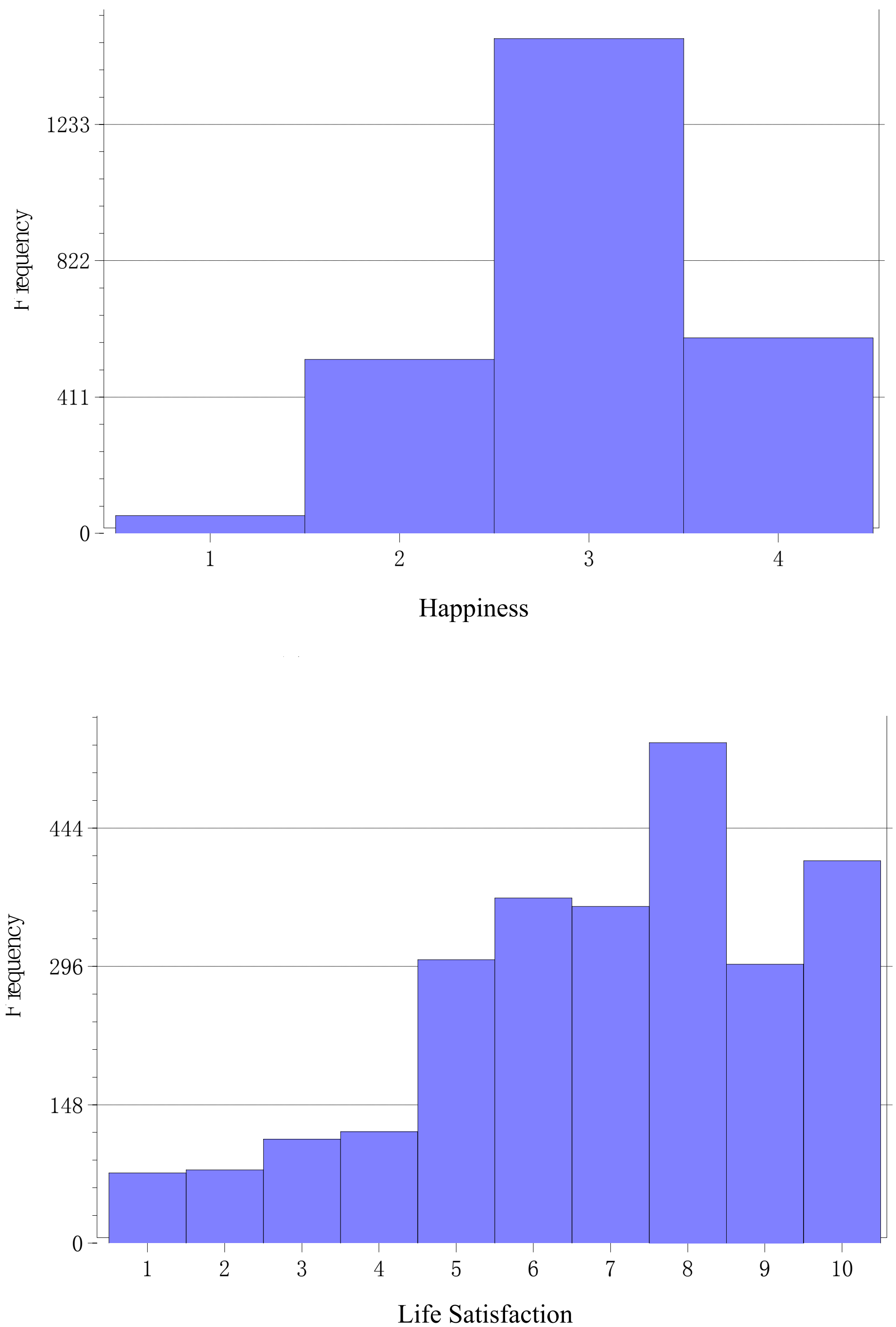
Table 6-1 Cross-tabulations between Happiness and Life Satisfaction

	Life Sat=1	Life Sat=2	Life Sat=3	Life Sat=4	Life Sat=5	Life Sat=6	Life Sat=7	Life Sat=8	Life Sat=9	Life Sat=10	Total
Hap=1	13	13	2	3	6	2	3	6	3	2	53
Hap=2	33	43	44	44	68	90	60	64	37	41	524
Hap=3	21	18	50	60	195	212	235	332	173	195	1491
Hap=4	8	4	15	12	34	65	62	133	85	171	589
Total	75	78	111	119	303	369	360	535	298	409	2657

Source: World Values Surveys (1990, 1995-1997, 1999-2000).

[107] The survey in year 1999-2000 covers 24 provinces, in year 1991 covers 11 regions without clear indication of which provinces, and in year 1995-1997 covers 6 main parts of China, still, without clear indication of which exact provinces.

Figure 6-1 Histograms of Happiness and Life Satisfaction



Source: World Values Surveys (1990, 1995-1997, 1999-2000)

The SES included in our analysis is composed of the education level, main occupation as a farmer, financial status and unemployment. Education level is captured by ages at which individuals finished full-time education, and the values are restricted from 8 to 30 in our estimate[108]. Financial status takes two forms: one is the objective gross household income, i.e. all income included last year, and the other is the subjective household financial satisfaction. We assume that the latter is picking up a comparison effect with their peers when respondents answer this question. Both total household income and household financial satisfaction are reported on a scale of 10 categories with ascending order, and on average, total household income is around 3000-4000 yuan per year.

In addition, self-reported health and choice or control of one's own life are also included in the estimate. Possible responses for self-reported health are: i) very good, ii) good, iii) fair and iv) poor[109]. The choice or control question is based on the questions of "Some people feel they have completely free choice and control over their lives, while others feel that what they do has no real effect on what happens to them. Please use the scale where '1' means 'none at all' and '10' means 'a great deal' to indicate how much choice and control you feel you have over the way your life turns out.". There is a relatively high level of choice or control of life with an average of '7' out of 10 points.

---

[108] Basic formal education includes primary education (normally six years), academic secondary education (normally three years of lower and three years of upper middle school) and academic higher education (normally four years of undergraduate degree, three years of master degree and three years of PhD).

[109] We have also reversed the scale of the health measure in the regression.

Table 6-2 Sample Size and Distribution of Happiness and Life Satisfaction (1990, 1995-1997, 1999-2000)

Variables	Whole Sample (%)	Male (%)	Female (%)
Dependent Variables			
Happiness			
Not at all Happy ('1')	1.99	1.74	2.32
Not very Happy ('2')	19.72	22.03	16.64
Quite Happy ('3')	56.12	55.40	57.03
Very Happy ('4')	22.17	20.72	24.01
Life Satisfaction			
Dissatisfied ('1')	2.82	3.15	2.40
'2'	2.94	2.88	3.00
'3'	4.18	4.49	3.77
'4'	4.48	4.83	4.03
'5'	11.40	11.07	11.84
'6'	13.89	14.08	13.64
'7'	13.55	14.08	12.86
'8'	20.14	18.78	21.87
'9'	11.22	10.66	11.92
Satisfied ('10')	15.39	15.96	14.67
Other Variables			
Female ('0')	43.88	-	-
Age	18 - 65	18 - 65	18 - 65
Married (Dummy)	81.78	82.23	81.22
Education - 'Age Left School'	8 - 30	8 - 30	8 - 30
Income			
Less than 500 Yuan ('1')	4.93	5.30	4.46
501-1000 yuan ('2')	12.46	13.41	11.23
1001-2000 yuan ('3')	16.75	17.84	15.35
2001-3000 yuan ('4')	12.95	13.48	12.26
3001-4000 yuan ('5')	17.43	15.63	19.73
4001-6000 yuan ('6')	13.59	13.41	13.81
6001-10000 yuan ('7')	10.99	10.60	11.49
10001-20000 yuan ('8')	7.72	7.04	8.58
20001-50000 yuan ('9')	2.78	2.55	2.40
50001-100000yuan ('10') and more	0.72	0.74	0.69
Unemployed (Dummy)	13.06	9.66	17.41
Farmer (Dummy)	36.88	35.35	38.85

Table 6-2 Continued

Variables	Whole Sample (%)	Male (%)	Female (%)
Financial Satisfaction with Household			
Dissatisfied ('1')	6.03	6.58	5.33
'2'	4.22	4.90	3.35
'3'	7.65	8.33	6.79
'4'	6.64	6.78	6.45
'5'	16.78	17.26	16.17
'6'	14.18	14.57	13.67
'7'	13.39	12.89	14.02
'8'	14.56	13.16	16.34
'9'	6.15	5.64	6.79
Satisfied ('10')	10.41	9.87	11.09
Self-reported Health			
Poor ('1')	6.25	4.83	8.06
Fair ('2')	29.36	28.17	30.87
Good ('3')	29.96	31.12	28.47
Very Good ('4')	34.44	35.88	32.59
Choice/Control of People's Own Lives			
No Choice/Control of life ('1')	3.42	3.22	3.69
'2'	2.56	2.35	2.83
'3'	3.76	3.22	4.46
'4'	3.84	3.82	3.86
'5'	11.93	11.60	12.35
'6'	12.19	11.60	12.95
'7'	13.81	13.75	13.89
'8'	18.67	19.18	18.01
'9'	9.41	9.59	9.18
A great deal of Choice/Control ('10')	20.40	21.66	18.78
Social Capital variables			
Participation in '0' Vol. Org. or Activities	76.67	76.53	76.84
'1'	16.18	16.16	16.21
'2'	5.01	5.30	4.63
'3'	1.73	1.61	1.89
'4'	0.41	0.40	0.43
Party Membership (Dummy)	13.81	16.50	10.38
Most people can be trusted (Dummy)	54.99	55.67	54.12
No. of the Observations	2657	1491	1166

Source: World Values Surveys (1990, 1995-1997, 1999-2000)

We also include data of social capital variables based on the individual level. Two participation variables separately relate to each individual's own participation in voluntary organizations and his/her party membership[110]. Party membership is not voluntary in the Chinese context, and here is mainly Communist party membership[111]. Given the data availability, participating voluntary organizations and activities included in our estimate are: conservation, environment, animal rights groups, sports or recreation, arts, music or cultural activities and other voluntary organizations and activities. We aggregate these four to one variable which measures the total number of organizations participating by each respondent. In addition, social trust is a binary variable included in our estimate, and it takes a value of '1' if respondents think that generally speaking, most people can be trusted. This relates to, as argued by Helliwell (2001), whether people think other people are trustworthy, rather than whether the respondents are trusting.

---

[110] Note that both the questions of participation in voluntary organizations and political party membership were asked differently in the third wave compared to the former two waves. In the third wave, these variables are binary variables taking a value of '1' if individual belongs to the associations and political parties. In the former two waves, however, it responds to three possibilities: i) belong to, ii) an active member and iii) not belong to. Therefore, to make three waves consistent, we recode them into binary variables based on the specifications of these two variables respectively. For participating voluntary organizations, we recode i) and ii) to take a value of '1', and iii) to take a value of '0'; for political party membership, we recode i) to take a value of '1', ii) and iii) to take a value of '0'.

[111] We have contacted one of the data-collecting staff by email, who was working in the Research Center for Contemporary China, Peking University. She replied our email and said that as far as they are aware, the party membership here is mainly Communist party membership.

6.4. Empirical Results and Analysis

6.4.1. The Impact of SES, Social Capital on SWB

Since the econometric framework of ordered probit model has already been discussed in the last chapter, we will directly start presenting the empirical results. Table 6-3 shows the estimated coefficients in separate equations of happiness and life satisfaction. The estimation was done in LIMDEP and checked in STATA. Beginning with the impact of gender, we can see that the coefficients are significantly negative for both equations, which indicates that females are happier and more satisfied than males. This is consistent with the results in the literature. From the statistical perspective and especially for happiness as shown in Table 6-2, we see that the percentage reporting “very happy” for females and males is 24.0% and 20.7% respectively. This is consistent with the regression results.

What about age? Our results show a strong U-shape pattern for life satisfaction, but not for happiness. The lowest point for life satisfaction is age 36 as shown in Figure 6-2. After reaching this point, life satisfaction rises significantly, with those around 45-50 as happy as those around 21-25, and those age 60 and up much happier still. This U-shape pattern for age is consistent with Blanchflower and Oswald’s (2004) finding for British and the United States. However, the lowest age 36 is less than the study by Sanfey and

Teksoz (2006), who examined SWB in transition countries and obtained the result that happiness declines with age until the early-fifties.

Figure 6-2 Age Impact on Life Satisfaction

Source: From Table 6-3.

Another variable which has a different impact for happiness and life satisfaction is the education level, captured by 'age left school'. 'Age left school' has a significant positive impact on the equation of happiness, but not on life satisfaction. This may be not too surprising. As mentioned before, life satisfaction may be more reflective of the long-term circumstance compared to happiness. Thus from a long-term perspective the beneficial effects of education may flow though higher incomes, better health, and higher perceived trust levels which have been already taken into account (see, e.g., Helliwell, 2003). Note that if we exclude these variables, and run a separate regression with the inclusion of gender, age, and year dummies, the estimated coefficient of 'age left school' changes to become significantly positive at 5% level. We do not report the estimated coefficients in the Table 6-3. As to the significant positive impact of 'age left school' on the equation relating to happiness, it may reflect a direct impact on happiness, which is consistent with the view in Scitovsky's (1976) joyless economy, that education facilitates abilities for people to enjoy life, through a great ability to enjoy paintings and culture for instance.

As in other research, marriage appears to be very conductive to SWB. Our results in Table 6-3 show that married people are more satisfied with lives than those who were never married, divorced, separated or widowed, and this positive linkage is more significant with respect to the equation of life satisfaction than to happiness. Following the early reviewed literature, we construct an interactive variable Married*Young (i.e., the 18-30 age group) to explore the potential age difference in the impact of being married on SWB, and we assume these are couples being newly-married. The reason we chose this age cohort is because we expect that many unstable factors which related to young couples may influence the results. The second and the fifth columns of Table 6-3 show that the coefficients of these interactive variables are significantly negative for both happiness and life satisfaction. Married people are happier and more satisfied than the non-married, but not for young couples. This may suggest the difficulties newly-marrieds face in both adapting to new lives and setting up home on their own. To an extent, the differential impact of being married and being newly-married confirms the causality running from marital status to SWB. In addition, it is also not clear, as some studies have argued, whether men or women benefit more from marriage. Separating married effect by gender showed no significant difference between men and women, and we do not report the estimated coefficients here.

As to the impact of economic factors on SWB, the unemployed are much less satisfied with their lives than others, and this linkage is only significant with respect to the equation of life satisfaction, but not to happiness. We lack the information on the regional rate of unemployment as well as how long the respondents have been unemployed, thus we can not comment on whether it is possible that people in regions with high unemployment feel less distressed than those living in areas with low joblessness. Note too that we also lack the information of respondents being in rural or urban areas, however, our results show that being a farmer is significantly negatively related to less happiness and lower levels of life satisfaction. To some extent, as the survey has defined the farmer as agricultural workers, the results capture some rural-urban differences on SWB.

Table 6-3 Determinants of Happiness and Life Satisfaction

Explanatory Variables	Life Satisfaction			Happiness		
Gender	-0.151***	-0.151***	-0.135***	-0.185***	-0.186***	-0.144***
	(-3.645)	(-3.653)	(-3.243)	(-4.097)	(-4.114)	(-3.166)
Age	-0.046***	-0.062***	-0.034***	-0.008	-0.036***	0.002
	(-3.304)	(-3.847)	(-2.422)	(-0.520)	(-2.084)	(0.122)
Agesq	0.0006***	0.0008***	0.0005***	0.0001	0.0004**	-0.00001
	(3.778)	(4.261)	(2.902)	(0.635)	(1.998)	(-0.080)
Married	0.235***	0.333***	0.196***	0.105*	0.275***	0.076
	(3.461)	(3.959)	(2.854)	(1.921)	(3.015)	(1.029)
Newly-married	-	-0.151**	-	-	-0.264***	-
(Married*age less than 30)		(-1.970)			(-3.171)	
Education-'Age Left School'	0.001	0.0002	-0.0002	0.016**	0.015*	0.018**
	(0.111)	(0.026)	(-0.022)	(2.072)	(1.936)	(2.263)
Total Household Income	0.045***	0.045***	-	0.039***	0.040***	-
	(4.578)	(4.611)		(3.222)	(3.293)	
Household Financial	-	-	0.281***	-	-	0.097***
Satisfaction			(30.951)			(9.983)
Unemployed	-0.218***	-0.233***	-0.107	-0.074	-0.098	-0.050
	(-3.309)	(-3.504)	(-1.593)	(-1.026)	(-1.348)	(-0.687)
Farmer	-0.177***	-0.179***	-0.106**	-0.115**	-0.116**	-0.123**
	(-3.806)	(-3.847)	(-2.277)	(-1.995)	(-2.024)	(-2.144)

Table 6-3 continued

Explanatory Variables	Life Satisfaction			Happiness		
Self-reported Health	0.244***	0.244***	0.185***	0.236***	0.237***	0.213***
	(10.924)	(10.933)	(8.153)	(9.625)	(9.650)	(8.638)
Choice/Control of One's	0.158***	0.157***	0.097***	0.071***	0.071***	0.043***
Life	(18.425)	(18.410)	(10.812)	(7.707)	(7.669)	(4.421)
Social Trust	0.133***	0.130***	0.124***	0.089**	0.084*	0.082*
	(3.283)	(3.211)	(3.040)	(2.025)	(1.907)	(1.861)
Participation in Voluntary	0.049*	0.049*	-0.005	0.085**	0.085**	0.073**
Organizations	(1.654)	(1.651)	(-0.158)	(2.508)	(2.516)	(2.153)
Party Membership	0.178***	0.173***	0.168***	-0.003	-0.012	0.0004
	(2.904)	(2.830)	(2.713)	(-0.039)	(-0.174)	(0.006)
Year 1990	0.652***	0.654***	0.283***	0.137*	0.140*	-0.038
	(9.146)	(9.178)	(4.513)	(1.773)	(1.813)	(-0.555)
Year 1995-1997	0.314***	0.314***	0.084	0.278***	0.277***	0.184***
	(5.831)	(5.818)	(1.572)	(4.728)	(4.709)	(3.194)
Log Likelihood	-5330.809	-5328.870	-4899.277	-2664.906	-2659.881	-2613.365
Restricted Log Likelihood	-5645.513	-5645.513	-5633.728	-2806.941	-2806.941	-2801.262
No. of the Observations	2657	2657	2652	2657	2657	2652

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions are estimated by ordered probit.

The idea that income buys happiness is one of the assumptions in microeconomics, see any textbook, and made for deductive reasons[112]. Studies on the effects of income on SWB provide an ambiguous picture and the main debate is based on the role of relative rather than absolute income in determining SWB. We include both objective household income and subjective household financial satisfaction in the estimate. Table 6-3 shows that both objective and subjective measures are significantly positive in impacting on SWB, although the latter is more significant than the former. This is consistent with others' results which emphasize the importance of relative income. In addition, the magnitude of the coefficients indicates that the linkage is stronger for life satisfaction than for happiness. Note that the significance of the other variables is largely as before in the equation for life satisfaction when we include subjective financial satisfaction instead of objective household income, however, the impact of unemployed status fails to be significant as shown in the third column of Table 6-3. As we assume that the subjective financial satisfaction is picking up a comparison effect with their peers, this indicates that from a long run perspective, the impact of unemployed status on life satisfaction is not just simply due to the loss of income, but, following the literature, to a large extent related to the comparison with others or an individual's aspirational standards of living.

Self-reported health and choice or control of one's own life are two particularly

[112] Blanchflower and Oswald (2004) argued that an indirect utility function is of course increasing in income, and consumer theory can be done using revealed preference alone.

significant determinants of happiness and life satisfaction. This is again consistent with others' results. People with very good health are happier and more satisfied with their lives, as are people with a great deal of choice or control which, as suggested, is closely related to people's choice of health behaviour (see, e.g., Bobak, *et al.*, 1998). Shields and Price (2005) argued that lower levels of individual well-being are correlated with different measures of poor health or disability, thus we may obtain more reliable results if we could include other indicators of health as a robustness check. As previously mentioned in the literature, the strong correlation between health and SWB exists for self-reported health measures, but not for objective health ratings by physicians[113]. However, individuals with a severe disabling condition or multiple or chronic health problems may report low SWB[114] (see e.g. Mehnert, *et al.*, 1990; Shields and Price, 2005). In this sense, it is worthwhile to examine whether the findings from the use of self-reported health still hold when some objective controls, such as acute/chronic physical illness, are included in the SWB estimate. However, as with many other studies the data was simply not available to allow us to do this. Finally there may also exist the possibility that these health-related coefficients may overstate the size of long-term effects of health status on SWB, since they may be both affected in the same direction by personality differences.

---

[113] Perceptions of health appear to be more important than objective health in their effects on SWB, although it is reasonable to assume that there is a linkage between these two.

[114] Several studies find evidence of partial adaptation to these circumstances over time. In addition, short-term effects may be larger than long-term outcomes as people adapt to their new health status and find new peer groups (see e.g. Kahneman, *et al.*, 1999).

Earlier research has found a strong correlation between an individual's participation in voluntary organizations, trust levels and SWB. There are two participation variables included in the estimates as mentioned in the data section. One variable is related to the participation in voluntary organizations, which is the typical variable included in this context based on the individual level; the other is related to an individual's party membership. However, since the party membership is not voluntarily participated in China, its interpretation is rather complex. The Chinese Communist party is the ruling political party of the China, and its position is guaranteed by the country's constitution[115]. For the individual, joining a party in the West may be seen as primarily marking an ideological affiliation, however, Communist party membership in China is likely to have a direct effect on one's career or economic circumstances. Thus in this sense, the Communist party membership as we assume is one type of 'social network', but not a typical indication of social capital.

Table 6-3 shows the results. The estimated coefficients for respondents' trust level are significantly positive for both the equation of happiness and life satisfaction, which indicates that those who think most people can be trusted are happier and more satisfied with their lives. As to the two participation variables, results are different for happiness

[115] There are eight 'democratic' parties in China, but these are relatively small and, as virtual satellite organizations of the Chinese Communist party, do not provide genuine political competition. Historically, these parties have tended to recruit people from non-labouring classes who are nonetheless supportive of the Chinese Communist party. Detailed discussions see Appleton, *et al.* (2005).

and life satisfaction. Individuals who are involved in more voluntary organizations have higher levels of happiness, and those who are a member of Communist party are more satisfied with their lives. Following the literature, the most convincing evidence of positive impact of social involvement on SWB is captured through health outcomes. In part the impact of participation in voluntary organizations on life satisfaction may flow though better health and control which have already included in the estimate.

If we exclude health status and control variable[116] and run a separate regression for both life satisfaction and happiness, results shown in Table 6-4 indicate that both participation variables are significantly related to life satisfaction, and the results for happiness remain unchanged. To put this differently, the significantly positive impact of party membership on life satisfaction does, apart from flowing through health outcomes, appear to go beyond the impact of general social involvement. This is consistent with our assumption and the fact the Chinese Communist party has a monopoly on political power and historically has maintained its grip on power. In addition, as with a direct impact of education on happiness, voluntary involvement also facilitates the ability and opportunity for people to enjoy life.

[116] Note that we start by excluding self-reported health status only, and the coefficient of participation in voluntary organizations is not significant. Similarly, when we exclude the control variable only, the coefficient is significant at 10% level. This indicates that the impact of social involvement on SWB flows through both health status and related control variable which is associated with people's choice of health behaviour.

Table 6-4 Determinants of Happiness and Life Satisfaction – excluding Health Status and Control Variable

Explanatory Variables	Life Satisfaction	Happiness
Gender	-0.064	-0.121***
	(-1.559)	(-2.705)
Age	-0.059***	-0.019
	(-4.300)	(-1.287)
Agesq	0.0007***	0.0002
	(4.374)	(1.099)
Married	0.291***	0.154**
	(4.306)	(2.105)
Education-'Age Left School'	-0.001	0.015*
	(-0.113)	(1.869)
Total Household Income	0.070***	0.063***
	(7.319)	(5.337)
Unemployed	-0.273***	-0.110
	(-4.164)	(-1.525)
Farmer	-0.139***	-0.074
	(-3.012)	(-1.304)
Social Trust	0.147***	0.104**
	(3.650)	(2.375)
Participation in Voluntary	0.073**	0.102***
Organizations	(2.480)	(3.043)
Party Membership	0.240***	0.037
	(3.954)	(0.557)
Year 1990	0.667***	0.200***
	(9.424)	(2.628)
Year 1995-1997	0.303***	0.308***
	(5.688)	(5.311)
Log Likelihood	-5574.570	-2752.435
Restricted Log Likelihood	-5645.513	-2806.941
No. of the Observations	2657	2657

Notes: t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. Regressions are estimated by ordered probit.

6.4.2. The Determinants of Social Capital

In this section we turn to an analysis of the determinants of social capital variables including participation in voluntary associations, party membership and social trust. We estimate the former by the ordered probit model and the latter two by the probit model based on the specifications of their dependant variables. The regression results are shown in Table 6-5. Unquestionably, the most significant impact on the social capital variables across individuals is the education level. This is consistent with the literature, and it is a powerful result, as we have estimated the impact of social capital on well-being, the total benefit of investment in education has been broadened.

We include age and age squared in the first column of Table 6-5 where these two variables are significant, and we simply report the results of including the log of age in the latter two columns. There is a significant nonlinearity for participation in voluntary associations, it first declines and then increases with a turning point of year 46. Note that this is not consistent with Glaeser's (2001) argument as he presented an inverted U-shape relationship between age and social capital. However, this fits China, and other developing countries' situations as people are fighting much harder to survive at a reasonable level when they are younger, and it is reasonable for them to participate in more voluntary organizations as they age. In addition, social trust, and specifically party membership significantly increase with age.

Table 6-5 The Determinants of Social Capital Variables

Explanatory Variables	Voluntary Organizations	Party Membership	Social Trust
Gender	0.009	0.167**	-0.028
	(0.154)	(2.385)	(-0.550)
Age or Log(age)	-0.037**	0.672***	0.156*
	(-1.994)	(5.365)	(1.685)
Agesq	0.0004*	-	-
	(1.784)		
Married	-0.143	0.076	0.062
	(-1.635)	(0.678)	(0.805)
Education-'Age Left	0.052***	0.090***	0.046***
School'	(5.719)	(7.634)	(5.202)
Total Household Income	-0.003	0.055***	0.012
	(-0.209)	(2.923)	(0.895)
Unemployed	-0.035	-0.190*	-0.144*
	(-0.401)	(-1.684)	(-1.857)
Farmer	-0.224***	-0.111	0.051
	(-3.201)	(-1.124)	(0.794)
Self-reported Health	0.045	0.016	0.030
	(1.471)	(0.431)	(1.079)
Choice/Control of One's	0.019	0.048***	0.014
Life	(1.628)	(3.129)	(1.314)
Year 1990	0.438***	0.736***	0.185**
	(4.075)	(6.308)	(2.150)
Year 1995-1997	1.123***	-0.145	-0.057
	(12.718)	(-1.502)	(-0.883)
Log Likelihood	-1792.948	-897.495	-1797.806
Restricted Log Likelihood	-1969.603	-1066.909	-1828.455
No. of the Observations	2657	2657	2657

Notes: When agesq is not reported age represents the log of age. t statistics in parentheses, *** Significant at 1%; **Significant at 5%; * Significant at 10%. The first column is estimated by ordered probit, the second and the third columns are estimated by probit.

Two economic factors need to be mentioned here: being a farmer is negatively linked with fewer involvements in voluntary associations, and total household income is positively related to the possibility of being party members. Both of them are significant at 1% level. The former is not too surprising as more generally social involvement or connectedness is restricted by occupational characteristics[117] and location; and the latter is to some extent consistent with our assumption that participation in political parties goes beyond simple social involvement. Finally, as shown by year dummies, the participation rates and trust levels have declined specifically compared to the year 1990.

6.5. Conclusions

In this chapter we have examined the extent to which social capital and SES impact upon subjective well-being using three waves of the World Value Survey covering 1990, 1995-1997 and 1999-2000. This is the first time this issue, specifically the impact of social capital, has been analyzed in mainland China. We included both life satisfaction and happiness in our analysis, and our results show similar results between them. But there are differences, in particular and as expected, we found socio-economic variables to be more significant in explaining life satisfaction compared to happiness. In this sense, we have confirmed the argument in the literature that happiness is more related to the short-term circumstances compared to the life satisfaction. Noticeable too in this context is the significance of education in explaining happiness, but not life satisfaction. This

[117] Individuals who work in occupations with more social contact are more likely to invest more in social connections.

impact is in addition to any impact of education on income or health which is separately included in the regression. This supports the view put forward by Scitovsky that education facilitates the enjoyment of cultural activities which can lead to short-term happiness, e.g. when reading a book or going to the theatre.

As with many other countries, we found a U-shaped relationship with respect to age and life satisfaction. Females are happier and more satisfied with lives, so are married people, but not young married people. We expanded the research to include a variable which reflects the degree of choice or control over people's lives and find it to have a positive impact. This variable which has been suggested to be closely related to people's choice of health behaviour, together with health status is the most significant of all the explanatory variables in determining SWB. Again as with others' results, we found the subjective measure of income, which as we assume is picking up a comparison effect with their peers, has more significant impact on happiness and more noticeable, life satisfaction.

We also found the impact of social capital on SWB. The impact of trust is consistent with findings of well-being research in general in which people who think most people can be trusted are happier and more satisfied with lives. We separated participation in voluntary associations and political party membership based on Chinese context and we found, as we expect, that individuals who are involved in more voluntary organizations have higher levels of happiness, and those who are a member of Communist party are more satisfied with their lives. Being a member of the Chinese Communist party is utterly different from being a member of any political party in the Western sense. The dominant position of the Chinese Communist party and its occupation of important administrative positions imply that the membership may bring additional benefits to the individual. According to Appleton, *et al.* (2005), the Chinese Communist party membership can be regarded as an investment of 'political capital' and subsequent benefits include additional income, additional perks, higher status and greater power or influence. In addition, as with others' results for different countries, education levels are the strongest and most consistent explanatory factors of both social and political involvement and trust levels. It is the key variable closely linked with both social capital and well-being. Thus it offers another route by which SWB can be enhanced.

Overall it is noticeable how similar in many, not all but many, respects our results are to those done in richer, market economies. This gives greater credence to this research in that the results appear valid in very differently cultural, economic and political contexts.

Chapter 7 Conclusion

China's transition from a planned to a market economy began at the end of 1978, and since the 1990s China has made rapid progress in economic growth. Economic reforms have brought significant changes in China's labour market and market determined wage rates. At the same time, higher education in China has experienced a massive expansion, especially over the last two decades, and the majority of high school graduates in China go on to attend college. Thus this reality has encouraged great interest in the new implications for the returns to education. The benchmark model for the development of empirical estimation of the returns to education has been developed by Mincer (1974), and a great deal of work based on this model has been done for many countries. Many studies have confirmed that education raises individual's productivity and subsequently future earnings. There has also developed the signaling/screening hypothesis of Arrow (1973) and Spence (1973, 1974) that employers use educational attainment to identify individuals with certain valuable innate traits that cannot be observed directly. Simply speaking, education acts as a signal of individuals' innate traits which enable employers to identify workers. This hypothesis has also been well-documented in the literature (see e.g., Brown and Sessions, 1998).

For decades, the primary argument for justifying education has been based on its direct impact on labour market productivity and earnings. Yet it is widely perceived that the effects of education spread beyond this to include social benefits for both individuals and societies, such as better health or higher levels of well-being. The health investment model developed by Grossman (1972) offers an important explanation that individuals inherit an initial stock of health that depreciates with age and can be increased by investment. Specifically, years of formal education completed is the most important correlate of good health and a healthy life style. Although there is a possibility that the direction of causality runs from better health to more education, or differences in one or more 'third variables' affect both education and health in the same direction. In addition, education is the strongest and most consistent explanatory factor for individuals with respect to their social capital, such as social connectedness and trust. A large number of empirical studies have reported a positive linkage between social capital and health, and subsequently well-being (see e.g., Putnam, 2000; Helliwell, 2001).

This study drew heavily on the work of these scholars, and our analysis expanded previous work in several important ways. We will discuss this in the next section of this chapter. Section 2 outlines the related policy implications, and in section 3, we will discuss the direction of further research.

7.1. Summary and Discussion

In this section, we structure the major findings of this thesis around six issues. *First, there is a noticeable increase in returns to education over time in China, and the return depends on gender and also indeed sectors and regions.* There exists an extensive literature on changes in returns to education associated with economic transition in other transitional economies. However, few studies have examined this issue in connection with the period of China's economic reform process. Previous studies have suggested that there are at least two competing hypotheses with respect to the returns to education during transition from centrally planned to market economy: one hypothesis is that labour market reforms at the onset would entail immediate increases in returns; the other is that the skills hitherto acquired can not be easily transferred to a changed economic situation, thus one would expect to see a temporary decline in returns during transition. Our results indicated that in 1989 the marginal return is 4%. This percentage declined from 1989 to 1993, and increased thereafter, specifically it has noticeably increased from 4% in 1997 to 7% in 2000. This is largely consistent with the latter hypothesis, although it is somewhat lower than some studies in both China and in other developing countries. We have discussed that one reason for our relatively low estimates is that the dependant variable includes subsidies and bonuses which tend to flatten the distribution. To a large extent, our estimates reflect the income (earnings and benefits) returns to education. As such they do not fully reflect productivity gains, but they do accurately reflect incentives.

Recent work in economic geography has emphasized the importance of proximity to large population mass in determining variables such as productivity. However, our results go further than this and emphasize the importance not simply of mass but proximity to the two large power centres in China, one economic, Shanghai, and the other political, Beijing. The impact of distance has similarities to a gravity model, as well as other spatial or geographic models. Our results indicated that people earn less, the further they are far from Beijing, and more noticeably, Shanghai; however, they do enjoy higher rate of returns to education in more remote regions. This suggests that the marginal productivity of labour and productivity is higher in Beijing and specifically Shanghai than other areas. We believe this is the first time this issue has been discovered in the Chinese context.

Second, we found strong evidence that education enhances ability and does not simply reflect it with the assumption that the value of education as a signal does not decline with age. It is open to debate in the literature whether education enhances productivity or just reflects it[118]. Researchers have also contested whether or not returns to education signals should decline as workers gain increased labour market experience. Followed the arguments of Riley (1979) and Farber and Gibbons (1996), a basic condition for a signaling equilibrium is that employers find out that their predictions based on the education signals are correct on average. We assume that on average the return to an educational signal is constant over time. The impact of education itself on wages may be because education either (i) *enhances* ability and productivity or (ii) because it *reflects* ability. We argued that when we have no direct measure of innate ability and simply include education on the right hand side it is picking up the influence of both innate ability and the impact of education in enhancing ability. We included an experience-education interaction variable in the earnings equation, and the impact of education on ability rather than simply proxying ability is confirmed by the existence of a significant negative sign on this variable. The implications are that education enhances ability, in part because of its impact on skills, then in a changing world as these skills date the

[118] As mentioned before, There exists the possibility that education both augments and signals individual productivity based on weak screening hypothesis as mentioned in chapter 3 (see e.g., Brown and Sessions, 1998). Thus our hypothesis (i) indicates the primary role of education is to augment marginal productivity.

impact of education on earnings should also decline. Note that the inclusion of interacted variables of education and experience has been implemented in the literature to test whether there is a ‘vintage effect’ of education (see, e.g., Liu, 1998), or whether there is greater depreciation in the return to experience for the better educated (see, e.g., Neuman and Weiss, 1995). However, our approach is based on this and further takes this declining rate of return to education over time as a sign for the impact of education in enhancing ability and productivity.

Third, a one-unit increase in education increases the probability of being in better health, and more educated people are also less likely to see a negative change in their health status. The observed correlation between education and health has been well-documented in the literature. An extensive review conducted by Grossman and Kaestner (1997) suggested that years of formal education completed is the most important correlate of good health. This is the health benefit of education which goes beyond the impact on individuals’ earnings, but perhaps too helps to explain that impact. In that a healthier worker is often a more productive one. There also exists the question of whether it is a unidirectional causal relation from education to health. There is a possibility that the direction of causality runs from better health to more education, or alternatively, differences in one or more ‘third variables’ which affect both education and health in the same direction. A number of studies have dealt with the endogeneity of education using instrumental variable techniques. The implications are to employ variables that are correlated with education but not correlated with such omitted third variables as ability, other inherited genetic traits, and time preference to obtain consistent estimates of education effects (Grossman, 2006). The difficulty here is that researchers must uncover instruments that plausibly are not correlated with third variables. In addition, if the instrument is weak, it may provide little or no useful information (Imbens and Rosenbaum, 2005). The problem of weak instruments is quite prevalent in recent IV studies, and researchers have pointed out that some IV studies have been undermined by a lack of precision in their first stage estimates (see, e.g., Bound, *et al.*, 1995).

Apart from looking at the direct effects of education on health, we also focused on whether the effect of education truly represents a causal effect. We focused on the hypothesis suggesting that the relationship between education and health can not be interpreted causally due to the unobserved inherent health. We deal with this in a slightly different way to others as we do not have suitably strong instruments in our dataset. We made the assumption that the changes in health[119] are independent of inherent health, i.e. that all people's health (stochastically) deteriorates at the same level given their educational attainment and possibly other variables too. We analyzed whether SES, especially education, has an impact on changes in health status. We create a dichotomous variable of health change, which takes a value of one if the individuals have a positive change or no change of their health status, and zero otherwise. Through this, we can find the SES pattern for people whose health statuses are getting worse. Note that our model is similar, but not the same as the traditional first-differences model. Our results indicated that education has a significantly positive impact on the change of health status, which means that individuals with more education are less likely to experience a negative change of health. Put literally education slows down the aging process.

[119] As mentioned, here we assume the answers of health question can be interpreted as cardinal.

Fourth, more household income significantly increases the probability of being in better health, and there is a significantly positive effect of income changes on an individual's health status. The relationship between income and health is one of the more heavily research topics in economics and the other social sciences. Greater income is associated with better health through, for example, a better lifestyle or diet, fewer monetary worries, better access to medical services and an improved living environment (Adler, *et al*., 1994; Smith, 1999). Similar to education, there are two main issues that need to be tackled in order to establish causality. These relate to: i) there are likely to be unobserved individual characteristics, which jointly determine both income and adult health; and ii) there exists the issue of reverse-causality. In order to establish the causal effect of income on health, recent studies have attempted to use more advanced panel data econometric techniques. For example, Frijters, *et al*. (2005) used the fixed-effects estimator, developed by Ferrer-i-Carbonell and Frijters (2004). Based on the GSOEP between 1984 and 2002, they concluded that there was evidence of a significant positive effect of income changes on health satisfaction, but the quantitative size of this effect is small. We applied the same methodology, estimating an ordered logit model with fixed-effects as well as an ordered probit model with random-effects. Our results indicated that the magnitude and significance of the coefficients in random-effects and fixed-effects models are very close. Thus the random-effects model has not been far away from the mark in the respect of unobservable heterogeneity.

Fifth, an individual's BMI should be adjusted for gender and age. We have identified the healthy BMI range after gender and age adjustment, and this healthy BMI has a *significant impact on health together with SES.* BMI is generally calculated as weight in kilograms over height in meters squared (weight(kg)/height(m)^2), and it is currently the most commonly used method in the empirical work. The WHO has devised a classification where by persons with BMIs below 19-25 are considered underweight, those with BMIs above this range are considered overweight or ‘at risk’ and those with BMIs greater than or equal to 30 are considered obese. The same healthy BMI is applicable for both men and women at different ages. Recent studies have indicated that BMI is age and gender dependent when used as an indicator of body fatness (see, e.g., Gallagher, *et al*., 1996). However, there are no clear cut-off points after adjusting for age and gender in the literature. Our analysis explored this by taking Log(weight) as a dependant variable and running the regression on Log(height), gender and age. Our results indicated that the power of height (in weight(kg)/height(m)^2) increases when account is taken of gender and age. For a given height, weight and age, a BMI value for a man is lower than for a woman; and similarly, for a given height, weight and gender, an individual’s BMI value gradually increases with age, until after around age 55. It is necessary to mention when we take Log(weight) as a dependant variable and run the regression on Log(height) only, the power of height is close to the standard measure of 2.00. In addition, education, household income and certain types of occupation also have impact on people’s BMI values, however, we cannot say that the healthy values of BMI for more educated people, for example, should be lower or higher than less educated people, it depends upon their other physical characteristics.

So far we have developed a little further the issue on how to define the BMI. We further identified the healthy BMI range based on this adjusted measure of the BMI. We assumed two critical values of α_L and α_H, and the range α_L - α_H is what we term the 'healthy range'. People outside this range are unhealthy, and the greater the distance the less healthy they are. We identify the critical points α_L and α_H through regression analysis. Self-reported health is regressed on (α_L – BMI), operative if BMI<$\alpha_{L,}$ and (BMI – α_H), operative if BMI>α_H, by an iterative search technique where the critical values for α_L and α_H included all possible combinations of individuals' BMI values. We chose the optimal combination on the basis of i) significantly negative coefficients for both (α_L – BMI) and (BMI – α_H), and ii) by identifying the best fit by the highest log likelihood ratio. Based on this, we identified the healthy range of 15 (α_L) to 19 (α_H).

We then constructed a new BMI variable based on the optimal value range of 15 (α_L) and 19 (α_H), defined as the sum of (15 – BMI), operative if BMI<15, and (BMI – 19), operative if BMI>19. The implications are the distance between individuals' BMI values and either 15 (α_L) or 19 (α_H) represents a deviation of BMI from its healthy range. Based on self-reported measure of health, our results confirmed that when a person's BMI value moves away from either 15 (α_L) or 19 (α_H), he/she tends to be less healthy. More importantly, we also found that the impact of this new BMI variable on health is affected by different SES. For a given BMI, there is a greater adverse impact on health for a man than a woman. Similarly, it also depends on age, and more specifically, it is convex, first increasing and then decreasing after reaching a certain age. In addition, higher income neutralizes to some extent an adverse new BMI. Finally and interestingly, a given value of new BMI has a greater adverse impact on health with more education. We have concluded that educated people tend to be healthier than less educated people, presumably because they have increased awareness of how to live a healthy life. An adverse BMI with more education on health, however, suggests that they are not utilizing this knowledge, thus reducing the 'educational advantage' in health.

Sixth, social capital has an impact on subjective well-being, and a key variable closely linked with both social capital and well-being is education. Most of the relevant work has been conducted in the Western countries, particular in the United States, and comparatively few studies have been done in other contexts. Because culture plays a potentially important role in this area, the relatively few studies based on Chinese participants is an obvious limitation. We identified three social capital variables based on the individual level: participation in voluntary organizations, Communist party membership and trust levels. We did not combine the former two based on the reasons that the party membership is not voluntary in the Chinese context. The Chinese Communist party is the ruling political party of the China, and being a Communist party member is likely to have a direct effect on one's career or economic circumstances. In this sense, it is one type of 'social network', but not a typical indication of social capital. According to Appleton, *et al.* (2005), the Chinese Communist party membership can be regarded as an investment of 'political capital'. To measure the subjective well-being, our analysis separated happiness and life satisfaction, and our results indicated that these two participation variables have different impact on happiness and life satisfaction respectively. Individuals who are involved in more voluntary organizations have higher levels of happiness, and those who are a member of Communist party are more satisfied with their lives. The dominant position of the Chinese Communist party and its occupation of important administrative positions imply that the membership may bring benefits to the individual, including for example additional income and higher social status.

We have confirmed the argument in the literature that happiness is more responsive to the short-term circumstances compared to life satisfaction. Education level, measured by ‘age left school’, has a significant positive impact on happiness, but not on life satisfaction. We have argued, also tested by the regression results, that from a long-term perspective the beneficial effects of education may flow though higher incomes, better health, and higher perceived trust levels which have been already taken into account. In addition, the significant impact of education on happiness may reflect a direct impact on happiness, as suggested by Scitovsky’s (1976) joyless economy, that education facilitates abilities for people to enjoy life. Furthermore, the most significant impact on the social capital variables across individuals is the education level. This is a powerful result, as we have estimated the impact of social capital on well-being, the total benefits of investment in education has been broadened.

7.2. The Implications for Educational Policy

In concluding this thesis, we also look at the implications of our results. This is a difficult task due to the fact that establishing a link between research results and policy initiatives is not always straightforward. However, we believe that the implications of our results warrant some detailed consideration by policy-makers.

Increased returns to education suggests there is an increasing demand for higher educated workers, given that they benefit from a wage premium when compared to those less educated. This is consistent with the situation, at least to the extent, that the higher education attainment of the Chinese labour force is low by international standards. In this respect, education is still a valuable investment from the private point of view. Educational attainment of the Chinese population has increased substantially during the transitional period. However, the returns rate to education is still low in international terms, and the total employment population ratio is also relatively low. According to the Ministry of Personnel, at the end of 2000, the national secondary and higher education or professionals of all personnel accounted for 8.9 percent for the nation’s total employed population. This indicates that it is important to establish a unified labour market with fair competition and freedom of movement. This also illustrates the potentially important interactions between the supply of the educated and labour market outcomes. More specifically, the focus should be on which specific, and also levels of, education are in high demand and should therefore be prioritised in the allocation of public resources.

We have found some evidence suggesting that returns to education date since the individual left formal education. This may be more noticeable for a transitional economy like China where skills acquired under the old regime are sometimes of little use in the new environment, and specifically for older people. Thus it is important for the Chinese government to offer some re-training or further-training programmes. On the other hand, it is also important for prospective students to know how to improve themselves, especially from a younger age. In this respective, some resources should also be devoted to the development of generic skills such as analytic ability, problem solving, and similar ‘learning skills’.

The benefits of education are larger than just earnings. Our results indicated that a one-unit increase in education increases the probability of being in better health. If, as we argued, this impact reflects a causal effect on health, a shift in public expenditure from health care to education may be more effective in improving both the level of education and the health status of the population. This suggests that the monetary value of the benefits of investing in education, including the health benefits, may be much larger than the cost of the investment. According to Groot and Maassen van de Brink (2006), the individual is the main beneficiary of a reduction in morbidity and mortality because of the higher educational attainment. In addition, investments in health behaviour by higher educated people may create positive externalities. Higher educated people, for example, are more likely to take exercise and eat healthy food (less likely to be overweight), and their healthy behaviour may encourage others to adopt similar behaviour.

Importantly, even if a large part of the association between education and health is causal, as suggested by our results that more educated people are indeed less likely to experience a negative change of their health status, questions may still remain. Individuals may be unaware of the health benefits of education when they make their education decisions. It is also possible that they do not fully utilize the knowledge to implement the 'educational advantage' in health. This is suggested by the significance of the interaction term between education and BMI, which indicates an adverse BMI reduces the health advantage of education.

According to Cutler and Lleras-Muney (2006), education policies do have the potential to have a substantial effect on health. Although there may be some uncertainly about the full causal linkage, it is far better for policy-makers to put efforts on more educational subsidies as opposed to less. At the same time, it is also important to introduce public awareness programmes on 'healthy living' which may diffuse the educational advantage across a wider range of people.

7.3. Future Research

The years since 1990 have brought major changes in many dimensions of China's economy. The rate of return to education has noticeably increased from 1989 to 2000. Still, it is somewhat lower than some studies in China and studies in other developing countries. We have argued that this may in part reflect the underdevelopment in China's market economy in the transition period since economic reform is a continuous process and most policies take time to produce any changes. China's labour market has indeed undertaken rapid changes, specifically after China entered the World Trade Organization in 2001. Important new developments include the emergence of private business as a leading source of new jobs and increased job mobility, especially among well-educated workers (Rawski, 2005). Thus it would be desirable to update estimates of returns to education. Future updates based on more recent data could not only provide more robust estimates, but also provide evidence on whether the performance of China's labour market is improving over time.

Furthermore, different from Western populations, our estimates indicated that there is a positive association between household income and being overweight. Similarly, according to Wang, *et al*. (2007), the prevalence of overweight and obesity increased in all gender and age groups and in all geographic areas between 1992 and 2002. Over the past decade, together with China's impressive economic developments, people have also experienced many dramatic changes in their lifestyles related to an increase in family income and an increasing availability of food owing to advances in agriculture and increased global trade**.** This may have some new implications on the linkage between SES and overweight status. Updated results based on more recent data will also provide clearer information to contribute to the formulation of national strategies and programmes based on the obesity and obesity-related diseases. Further it is also important to know whether these results are replicated for other developing/transitional countries.

We have confirmed that the concept of the BMI is a valid one in the literature, and have further argued that it is reasonable that it should be calculated in such a manner as adjusts for gender and age. Note that we identified this adjusted BMI through purely an economic point of view, it may still worth further analysis by either health economists or anthropometrists. Based on the data availability when doing this research, we did not include the behavioural variables in our analysis for either identifying the healthy BMI range, or the impact of this healthy BMI on health. It is as suggested that the BMI, as an element of a life style, is closely related to health behaviour. Thus more robust results would be obtained if one estimates the impact of BMI on health together with different health behaviour. For example, we concluded that the impact of BMI on health is affected by gender, age, education and income, and we expect it may also be affected by whether people are smoking, taking regular exercises, etc. Another concern is that we examined the healthy range of this adjusted BMI based on the self-reported measure of health. There is evidence in the literature suggesting that it is a powerful predictor of subsequent mortality (see e.g. Idler and Kasl, 1995; Idler and Benyamini, 1997), still, it may not be as accurate as the morbidity and mortality data. To put it differently, it will be of value to further confirm the healthy BMI range based on morbidity and mortality data, although care will need to be taken as the BMI may experience sharp changes in the months prior to death. Finally, for a comparison purpose, it is also important to examine the cut-off points based on this adjusted measure of BMI across different ethic groups. We believe this to be a generally new contribution to the literature, and one which has the potential to be built upon in other economies, including the developed ones such as in North America and Europe.

Appendix I Tobit Model

The Tobit[120] model is a well-known econometric regression model used in the presence of censored data. In econometric research, there are frequently subjects for whom we do not observe the true response or dependent variable, all that is known is that the true response, if it had been observed, would have been above, (or below) some threshold. The Tobit model is a special case of a censored regression model, as the index variable, sometimes called the latent variable, y_i^* cannot always be observed while the independent variable x_i is observable.

The general formulation of the Tobit model is usually given by:

$$y_i^* = x_i^{'}\beta + u_i, \qquad u_i | x_i \sim \mathrm{N}(0, \sigma^2)$$
$$y_i = 0 \qquad \text{if } y_i^* \leq 0,$$
$$y_i = y_i^* \qquad \text{if } y_i^* > 0. \tag{A.1}$$

where y_i^* is a latent variable generated by the classical linear regression model. x_i is a vector of regressors, possibly including 1 for the intercept, and β is the corresponding vector of parameters. The error term u_i is assumed to be independent $\mathrm{N}(0, \sigma^2)$ distributed. According to Wooldridge (2002), for the model (A.1) to make sense, the variable y_i^* should have characteristics of a normal random variable, which indicates that the variable

[120] Tobit estimation was originally developed by Tobin (1958).

of interest y_i^* should have a homoskedastic normal distribution. This applies to our analysis in the section 5.5.4, where y_i^* denotes the 'true' distance of individuals' BMI values from the healthy range (19-25). An individual with an observed value of 0 (i.e. his/her BMI value is within (19-25)) has a true $y_i^* \leq 0$. Thus the OLS estimation will produce biased and inconsistent estimate of β.

The log-likelihood for the Tobit model is[121]:

$$\ln L = \sum_{y_i > 0} -\frac{1}{2}\left[\log(2\pi) + \ln\sigma^2 + \frac{(y_i - x_i^{'}\beta)^2}{\sigma^2}\right] + \sum_{y_i = 0} \ln\left[1 - \Phi\left(\frac{x_i^{'}\beta}{\sigma}\right)\right] \tag{A.2}$$

The two parts in equation (A.2) correspond to the classical regression for the nonlimit observations and the relevant probabilities for the limit observations, respectively.

Further we fit the Tobit model with individual random-effects:

$$y_{it}^* = x_{it}^{'}\beta + u_{it}, \qquad i = 1,2,\ldots,N \qquad t = 1,2,\ldots,T$$

$$u_{it} = v_i + \varepsilon_{it} \qquad u_{it} \sim \mathrm{N}(0,\Sigma), \qquad \Sigma = \begin{bmatrix} \sigma_v^2 & \sigma_v\sigma_\varepsilon \\ & \sigma_\varepsilon^2 \end{bmatrix} \tag{A.3}$$

Similarly the observed variable is:

[121] For detailed discussions, see Greene (2003).

$y_{it} = 0 \quad \text{if } y_{it}^* \le 0,$

$$y_{it} = y_{it}^* \quad \text{if } y_{it}^* > 0. \tag{A.4}$$

The composite error term u_{it} is assumed to be normally distributed with the characteristics described by (A.3). It is composed by a time-invariant individual random effects, v_i, and a time-varying idiosyncratic random error, ε_{it}.

Empirically, the output (e.g. estimated by STATA) statistically includes the overall and panel-level variance components together with ρ calculated as follows:

$$\rho = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\varepsilon^2} \tag{A.5}$$

It is the percent contribution to the total variance of the panel-level variance component. When ρ is zero, the panel-level variance component is unimportant and the panel estimator is not different from the pooled estimator.

Appendix II Probit Model

A Probit model is defined by:

$$\Pr(Y=1 \mid x) = \Phi(x^{'}\beta) \qquad \text{(A.6)}$$

The function $\Phi(.)$ is a commonly used notation for the standard normal distribution. Other distribution functions can be used in place of Φ. The parameters β are typically estimated by maximum likelihood.

The Probit model can be generated by a latent variable model:

$$y_i^* = x_i^{'}\beta + u_i, \qquad u_i \mid x_i \sim \mathrm{N}(0,1) \qquad \text{(A.7)}$$

In this formulation, $x_i^{'}\beta$ is called the index function. Suppose that y_i is an indicator for whether the latent variable y_i^* is positive:

$$y_i = 1 \qquad \text{if } y_i^* > 0,$$
$$y_i = 0 \qquad \text{if } y_i^* \leq 0. \qquad \text{(A.8)}$$

This applies to our analysis in the section 5.5.3, where y_i^* denotes the change of individuals' health status. In our case, an individual with an observed value of 1 if the he/she has a positive change or no change of his/her health status, and 0 otherwise.

The log-likelihood for the Probit model is:

$$\ln L = \sum_{y_i=0} \ln[1 - \Phi(x_i^{'}\beta)] + \sum_{y_i=1} \ln \Phi(x_i^{'}\beta) \tag{A.9}$$

Further we fit the Probit model with individual random-effects:

$$\begin{aligned} &y_{it}^* = x_{it}^{'}\beta + u_{it}, \qquad i = 1,2,...,N \qquad t = 1,2,...,T \\ &u_{it} = v_i + \varepsilon_{it} \end{aligned} \tag{A.10}$$

and

$$\begin{aligned} &y_{it} = 1 \qquad \text{if } y_{it}^* > 0, \\ &y_{it} = 0 \qquad \text{if } y_{it}^* \leq 0. \end{aligned} \tag{A.11}$$

where y_{it}^* denotes the unobservable variable and y_{it} is the observed outcome, β is the vector of coefficients associated with the x_{it}, v_i denotes the individual specific effects and the ε_{it} is a random error. The underlying variance $\sigma^2 = \sigma_v^2 + \sigma_\varepsilon^2$ is not identified. We adopt the normalization $\sigma_\varepsilon^2 = 1$, and thus the overall error variance equals (σ_v^2+1).

Empirically, the standard deviation σ_v is also included in the output (e.g. estimated by STATA) together with ρ calculated as follows:

$$\rho = \frac{\sigma_v^2}{\sigma_v^2 + 1} \tag{A.12}$$

It is the proportion of the total variance contributed by the panel-level variance component. When ρ is zero, the panel-level variance component is unimportant and the panel estimator is not different from the pooled estimator.

Bibliography

Adams, S. J., 2001, "Educational Attainment and Health Evidence from a Sample of Older Adults", *Education Economics*, Vol. 10(1): 97-109.

Adams, P., Hurd, M., McFadden, D., Merrill, A. and Ribeiro, T., 2003, "Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status", *Journal of Econometrics*, Vol. 112(1): 3-56.

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L. and Syme, S. L., 1994, "Socioeconomic Status and Health, the Challenge of the Gradient", *American Psychologist*, Vol. 49(1): 15-24.

Al Riyami, A., Afifi, M. and Mabry, R. M, 2004, "Women's Autonomy, Education and Employment in Oman and their Influence on Contraceptive Use", *Reproductive Health Matters*, Vol. 12(23): 144-154.

Angrist, J. and Krueger, A., 1991, "Does Compulsory School Attendance Affect Schooling and Earnings?", *Quarterly Journal of Economics*, Vol. 106(4): 979-1014.

Appleton, S., Knight, J., Song, L. and Xia, Q. J., 2005, "The Economics of Communist Party Membership the Curious Case of the Rising Numbers and Wage Premium during China's Transition", Discussion Paper No. 2, China Policy Institute, University of Nottingham.

Arendt, J. N., 2005, “Does Education Cause Better Health? A Panel Data Analysis using School Reforms for Identification”, *Economics of Education Review*, Vol. 24(2): 149-160.

Argyle, M., 1987, *The Psychology of Happiness*. London: Methuen.

Arrazola, M., Hevia, J. D., Risueno, M. and Sanz, J. F., 2005, “A Proposal to Estimate Human Capital Depreciation: Some Evidence for Spain”, *Spanish Public Property/Magazine of Public Economy*, Vol. 172(1): 9-22.

Arrow, K., 1973, “Higher Education as a Filter”, *Journal of Public Economics*, Vol. 2(3): 193-216.

Ashenfelter, O. and Krueger, A. B., 1994, “Estimates of the Economic Returns to Schooling for a New Sample of Twins”, Princeton University Industrial Relations Section Working Paper, No. 304.

Ashenfelter, O. and Zimmerman, D., 1997, “Estimating of Return to Schooling from Sibling Data: Fathers, Sons and Brothers”, *Review of Economics and Statistics*, Vol. 79: 1-9.

Baltagi, B. H., 2005, *Econometric Analysis of Panel Data*, 3rd edition, John Wiley & Sons, NewYork.

Becker, G. S.， 1964， *Human Capital: A theoretical and Empirical Analysis, with Special Reference to Education*, Columbia University Press, New York.

Becker, G. S.，1964，*Human Capital*. 2nd edition, New York, NY: National Bureau of Economic Research, 1975.

Becker, G. S., 1994, *Human Capital*. 3rd edition, Chicago：The University of Chicago Press.

Behrman, J. R., 1997, “Concept and Measurement Issues”, in *The social benefits of education*, edited by Behrman, J. R. and Stacey, N., University of Michigan Press.

Behrman, J. R., Crawford, D. L. and Stacey, N., 1997, *The social benefits of education*, edited by Behrman, J. R. and Stacey, N., University of Michigan Press.

Behrman, J. R. and Wolfe, B. L., 1989, “Does More Schooling Make Women Better Nourished and Healthier? Adult Sibling Random and Fixed Effects Estimates for Nicaragua”, *The Journal of Human Resources*, Vol. 24(4): 644-663.

Bell, A. C., Adair, L. S. and Popkin, B. M., 2002, “Ethnic Differences in the Association between Body Mass Index and Hypertension”, *American Journal of Epidemiology*, Vol. 155(4): 346-353.

Belman, D. and Heywood, J. S., 1997, “Sheepskin Effects by Cohort: Implications of Job Matching in a Signaling Model”, *Oxford Economic Papers*, Vol. 49(4): 623-637.

Benjamin, D., Brandt L., Giles, J. and Wang, S. G., 2005, “Income Inequality during China’s Economic Transition”, Working Papers, No. 238, Department of Economics, University of Toronto.

Benzeval, M. and Judge, K., 2001, “Income and Health: The Time Dimension”, *Social Science and Medicine*, Vol. 53(9): 1371-90.

Benzeval, M., Taylor, J. and Judge, K., 2000, “Evidence on the Relationship between Low Income and Poor Health: Is the Government Doing Enough?”, *Fiscal Studies*, Vol. 21(3): 375-399.

Berger, M. C. and Leigh, J. P., 1989, “Schooling, Self-Selection, and Health”, *The Journal of Human Resources*, Vol. 24(3): 433-455.

Berkman, L. F. and Syme, S. L., 1979, “Social Networks, Host Resistance, and Mortality: A Nine-Year Follow-Up Study of Alameda County Residents”, *American Journal of Epidemiology*, Vol. 109: 186-204.

Beydoun, M. A. and Popkin, B. M., 2005, “The Impact of Socio-Economic Factors on Functional Status Decline among Community-Dwelling Older Adults in China”, *Social Science and Medicine*, Vol. 60: 2045-2057.

Blanchflower, D. G. and Oswald, A. J., 2000, “Well-being over Time in Britain and the USA”, NBER Working Paper, No. 7487. National Bureau of Economic Research, Cambridge.

Blanchflower, D. G. and Oswald, A. J., 2004, “Well-being over Time in Britain and the USA’, *Journal of Public Economics*, Vol. 88(7-8): 1359-1386.

Bobak, M., Hertzman, C., skodova, Z. and Marmot, M., 1998, “Association between Psychosocial Factors at Work and Nonfatal Myocardial Infarction in a Population-Based Case-Control Study in Czech Men”, Vol. 9(1): 43-47.

Böckerman, P. and Ilmakunnas, P., 2007, “Unemployment and Self-assessed Health: Evidence from Panel Data”, Discussion Paper, No. 148, Helsinki Centre of Economic Research, ISSN 1795-0562.

Borg, V. and Kristensen, T. S., 2000, “Social Class and Self-rated Health: Can the Gradient be Explained by Differences in Life Style or Work Environment?”, *Social Science and Medicine*, Vol. 51(7): 1019-1032.

Bortolotti, B., Fantini, M. and Sinisalco, D., 2004, “Privatisation around the World: Evidence from Panel Data”, *Journal of Public Economics*, Vol. 88(1-2): 305-322.

Bosma, H., Marmot, M. G., Hemingway, H., Nicholson, A. C., Brunner, E. and Stansfeld, S. A., 1997, “Low Job Control and Risk of Coronary Heart Disease in Whitehall II (prospective cohort) Study”, *British Medical Journal*, Vol. 314(7080): 558-565.

Bound, J., Jaeger, D. A. and Baker, R., 1995, “Problems with Instrumental Variables Estimation when the Correlation between the Instruments and the Endogenous Explanatory Variables is Weak”, *Journal of American Statistical Association*, Vol. 90(430): 443-450.

Brenner, J., 2006, “Parental Impact on Attitude Formation: A Sibling Study on Worries about Immigration”, MPRA paper, No. 34.

Breusch, T. S. and Pagan, A. R., 1980, “The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics”, *Review of Economic Studies*, Vol. 47(1): 239-53.

Brooks, R. and Tao, R., 2003, “China’s Labour Market Performance and Challenges”, IMF Working Paper, No. 03/210.

Brown, S. and Sessions, J. G., 1998, “Education, Employment Status and Earnings: A Comparative Test of the Strong Screening Hypothesis”, *Scottish Journal of Political Economy*, Vol. 45(5): 586-591.

Buch, C. M., Kleinert, M. and Toubal, F., 2004, “The Distance Puzzle: On the Interpretation of the Distance Coefficient in Gravity Equations”, *Economics Letters*, Vol. 83(3): 293-298.

Burkhauser, R. V. and Cawley, J., 2008, “Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research”, *Journal of Health Economics*, Vol. 27(2): 519-529.

Burstrom, B. and Fredlund, P., 2001, “Self-rated Health: Is It as Good A Predictor of Subsequent Mortality among Adults in Lower as well as in Higher Social Classes?”, *Journal of Epidemiology and Community Health*, Vol. 55: 836-840.

Butler, J. and Moffit, R., 1982, “A Computationally Efficient Quadrature for the One-factor Multinomial Probit Model”, *Econometrica*, Vol. 50: 761-764.

Byron, R. and Manaloto, E., 1990, "Returns to Education in China", *Economic Development and Cultural Change*, Vol. 38 (4): 783-796.

Cameron, A. C. and Trivedi, P. K., 2005, *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.

Campos, N. F. and Jolliffe, D., 2002, "After, Before, and During: Returns to Education in the Hungarian Transition", William Davidson Institute Working Papers Series, No. 475.

Card, D. E., 1994, "Earnings, Schooling, and Ability Revisited", NBER Working Paper series, No. 4832, National Bureau of Economic Research.

Card, D. E., 1999, "The Causal Effect of Education on Earnings", *Handbook of Labour Economics*, Vol. 3(A): 1801-1863, O. C. Ashenfelter and D. E. Card (eds), Amsterdam: North-Holland.

Card, D. E., 2000, "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", Working Paper, No. 7769, National Bureau of Economic Research, Cambridge.

Carliner, G., 1982, "The Wages of Older Men", *The Journal of Human Resources*, Vol. 17(1): 25-38.

Case, A., Fertig, A. and Paxson, C., 2003, "From Cradle to Grave? The Lasting Impact of Childhood Health and Circumstance", NBER Working Paper, No. W9788.

Case, A., Lubotsky, D. and Paxson, C., 2002, "Economic Status and Health in Childhood: The Origins of the Gradient", *American Economic Review*, Vol. 92(5): 1308-1334.

Chamberlain, G., 1980, "Analysis of Covariance with Qualitative Data", *Review of Economic Studies*, Vol. 74: 225-238.

Chan, K. W. and Hu Y., 2003, "Urbanization in China in the 1990s: New Definition, Different Series, and Revised Trends", *China Review*, Vol. 3(2): 49-71.

Chandrasekhar, C. P. and Ghosh J., 2006, "Rising Inequality in China", Financial Daily from the HINDU group of publications, the *HINDU Business Line*, Jan (03).

Chase, R. S., 1998, "Markets for Communist Human Capital: Returns to Education and Experience in the Czech Republic and Slovakia", Centre Discussion Paper, No. 770, Yale University.

Cheidvasser, S. and Benitez-Silva, H., 2000, "The Educated Russian's Curse: Returns to Education in the Russian Federation", Department of Economics Working Papers, SUNY-Stony Brook, No. 00-05.

Chen, Y., Demurger, S. and Fournier, M., 2005, "Wage Differentials and Ownership Structure in Chinese Enterprises", *Economic Development and Cultural Change*, Vol. 53: 933-958.

Cheung, J. C. K., Kwan, A. Y. H., Chan S. S. C., Ngan, R. M. H., Ng, S. H., Leung, E. M. F. and Lau, A., 2005, "Quality of Life in Older Adults: Benefits from Caring Services in Hong Kong", *Social Indicators Research*, Vol. 71(1-3).

Chevalier, A., Harmon, C., Walker, I. and Zhu, Y., 2004, “Does Education Raise Productivity or Merely Reflect It?”, *Economic Journal*, Vol. 114(499): F499-531.

Cho, In-K. and Kreps, D., 1987, “Signaling Games and Stable Equilibria”, *Quarterly Journal of Economics*, Vol. 102: 179-221.

Choo, V., 2002, “WHO Reassesses Appropriate Body-mass Index for Asian Populations”, *The Lancet*, Vol. 360(9328): 235.

Chow, G. C., 2006, “An Economic Analysis of Health Care in China”, CEPS Working Paper, No. 132.

Chow, I. Hau-Siu, 1995, “Career Aspirations, Attitudes and Experiences of Female Managers in Hong Kong”, *Women in Management Review*, Vol. 10(1): 28-32.

Clark, A. E. and Oswald, A. J., 1994, “Unhappiness and Unemployment”, *Economic Journal*, Vol. 104(424): 648-659.

Clark, A. E. and Oswald, A. J., 1996, “Satisfaction and Comparison Income”, *Journal of Public Economics*, Vol. 61(3): 359-381.

Cohn, E. and Geske, T., 1990, *The Economics of Education*, 3rd edition, Oxford, Pergamon.

Contoyannis, P. and Jones, A. M., 2004, **“**Socio-economic Status, Health and Lifestyle”, *Journal of Health Economics*, Vol. 23(5): 965-995.

Contoyannis, P., Jones, A. and Rice, N., 2004, "The Dynamics of Health in the British Household Panel Study", *Journal of Applied Econometrics*, Vol. 19(4): 473-503.

Crawford, P. B., Story, M., Wang, M. C., Ritchie, L. D. and Sabry, Z. I., 2001, "Childhood and Adolescent Obesity: Ethnic Issues in the Epidemiology of Childhood Obesity", *Pediatric Clinics of North America*, Vol. 48(4): 855-878.

Cruz, L. M. and Moreira, M. J., 2005, "On the Validity of Econometric Techniques with Weak Instruments", *Journal of Human Resources*, Vol. 40 (2): 393-410.

Cutler, D. M. and Lleras-Muney, A., 2006, "Education and Health: Evaluating Theories and Evidence", Working Paper, No. 12352, National Bureau of Economic Research.

D'Addio, A. C., Eriksson, T. and Frijters, P., 2003, "An Analysis of the Determinants of Job Satisfaction when Individuals' Baseline Satisfaction Levels may Differ", CAM Working Papers, 2003-16, University of Copenhagen.

De Grip, A. and Van Loo, J., 2002, "The Economics of Skills Obsolescence: A Review", *Research in Labour Economics*, Vol. 21: 1-25.

De Vasconcellos, M. T. L., 1994, "Body Mass Index: Its Relationship with Food Consumption and Socio-economic Variables in Brazil", *European Journal of Clinical Nutrition*, Vol. 48(3): 115-123.

Delpeuch, F., Cornu, A., Massamba, J. P., Traissac, P. and Maire, B., 1994, "Is Body Mass Index Sensitively Related to Socio-economic Status and to Economic Adjustment?

A Case Study from the Congo", *European Journal of Clinical Nutrition*, Vol. 48(3): 141-147.

Deolalikar, A. B., 1993, "Gender Differences in the Returns to Schooling and in School Enrollment Rates in Indonesia", *Journal of Human Resources*, Vol. 28(4): 899-932.

Dickson, B. J. and Rublee, M. R., 2000, "Membership has its Privileges: The Socioeconomic Characteristics of Communist Party Members in Urban China"**,** *Comparative Political Studies*, Vol. 33(1): 87-112.

Diener, Ed., Suh, E. M., Lucas, R. E. and Smith, H. L., 1999, "Subjective Well Being: Three Decades of Progress", *Psychological Bulleting*, Vol. 125(2): 276-302.

Ding, D. Z. and Warner, M., 2001, "China's Labour-Management System Reforms: Breaking the 'Three Old Irons' (1978-1999)", *Asia Pacific Journal of Management*, Vol. 18: 315-334.

Dockery, A. M., 2004, "Happiness, Life satisfaction and the Role of Work: Evidence from Two Australian Surveys", Working Paper, University of Melbourne.

Dong, W. Z., 2001/2, "Health Care Reform in Urban China", Working Paper, Munk Centre for International Studies at the University of Toronto, Comparative Program on Health and Society.

Donovan, N. and Halpern, D., 2002, "Life Satisfaction: The State of Knowledge and Implications for Government", Discussion Paper, Strategy Unit of the Cabinet Office, London.

Duetz, M. S., Abel, T., Meier, C., and Niemann, S., 2000, “Self-rated Health, Life Satisfaction and Personal Characteristics of Post-menopausal Women under Estrogen Replacement Therapy”, *The European Menopause Journal*, Vol. 35: 71-79.

Duraisamy, P., 2000, “Changes in Returns to Education in India, 1983-94: by Gender, Age-Cohort and Location”, Centre Discussion Paper, No. 815, University of Madras.

Easterlin, R. A., 1995, “Will Raising the Incomes of All Increase the Happiness of All?”, *Journal of Economic behaviour and organisation*, Vol. 27(1): 35-48.

Eddington, N. and Shuman, R., 2005, “Subjective Well-being (Happiness)”, *Continuing Psychology Education*.

Ettner, S., 1996, “New Evidence on the Relationship between Income and Health”, *Journal of Health Economics*, Vol. 15(1): 67-85.

Farber, H. S. and Gibbons, R., 1996, “Learning and Wage Dynamics”, *Quarterly Journal of Economics*, Vol. 111(4): 1007-1047.

Feng, X., Tang, S. L., Bloom, G., Segall, M. and Gu, X. Y., 1995, “Cooperative Medical Schemes in Contemporary Rural China”, *Social Science and Medicine*, Vol. 41(8): 1111-1118.

Ferrer-i-Carbonell, A. and Frijters, P., 2004, “How Important is Methodology for the Estimates of the Determinants of Happiness?”, *The Economic Journal*, Vol. 114(497): 641-659.

Ferrer-i-Carbonell, A., 2005, “Income and Well-being: An Empirical Analysis of the Comparison Income Effect”, *Journal of Public Economics*, Vol. 89: 997-1019.

Filer, R. K, Hamermesh, D. S. and Rees, A. E., 1996, *The Economics of Work and Pay*, 6th edition. New York: HarperCollins College Publishers.

Flanagan, R., 1998, “Were Communists Good Human Capitalists? The Case of the Czech Republic,” *Labour Economics*, Vol. 5(2): 295-312.

Frechette, G., 2001, “sg158: Random-effects Ordered Probit”, *Stata Technical Bulletin*, Vol. 59: 23-27.

Frey, B. S. and Stutzer, A., 2002, *Happiness and Economics*. *How the Economy and Institutions Affect Well-Being*. Princeton U. P., Princeton, NJ.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A., 2003, “Estimating The Causal Effect of Income on Health: Evidence from Post Reunification East Germany”, Discussion Papers, No. 465, Centre for Economic Policy Research, Research School of Social Sciences, Australian National University.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A., 2004, “Money Does Matter: Evidence from Increasing Real Incomes and Life Satisfaction in East Germany following Reunification”, the *American Economic Review*, Vol. 94: 730-740.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A., 2005, “The Causal Effect of Income on Health: Evidence from German Reunification”, *Journal of Health Economics*, Vol. 24(5): 997-1017.

Frumkin, H., 2002, “Urban Sprawl and Public Health”, *Public Health Reports,* Vol. 117: 201-217.

Fuchs, V. R., 1982, “Time Preference and Health: An Exploratory Study,” in Victor R. Fuchs ed. *Economic Aspects of Health*. Chicago: University of Chicago Press.

Fuchs, V. R., 1996, “Economics, Values, and Health Care Reform”, *The American Economic Review,* Vol. 86(1): 1-24.

Fuchs, V. R., 2004, “Reflections on the Socio-Economic Correlates of Health”, *Journal of Health Economics*, Vol. 23(4): 653-661.

Gallagher, D., Visser, M., Sepulveda, D., Pierson, R. N., Harris, T. and Heymsfield, S. B., 1996, “How Useful Is Body Mass Index for Comparison of Body Fatness across Age, Sex, and Ethnic Groups?”, *American Journal of Epidemiology,* Vol. 143(3): 228-239.

Gallo, L. C. and Matthews, K. A., 1999, “Do Negative Emotions Mediate the Association between Socioeconomic Status and Health?”, *Annals of the New York Academy of Sciences*, Vol. 896: 226-245.

Gardner, J. and Oswald, A., 2001, “Does Money Buy Happiness? A Longitudinal Study Using Data on Windfalls”, University of Warwick Mimeo.

Gibson, J., 2000, “Sheepskin Effects and the Returns to Education in New Zealand: Do They Differ by Ethnic Groups?’, *New Zealand Economic Papers,* Vol. 34(2): 201-220.

Gibson, J. and Fatai, O. K., 2006, “Subsidies, Selectivity and the Returns to Education in Urban Papua New Guinea”, *Economics of Education Review*, Vol. 2: 133-146.

Gilleskie, D. B. and Harrison, A. L., 1998, “The Effect of Endogenous Health Inputs on the Relationship between Health and Education”, *Economics of Education Review*, Vol. 17(3): 279-295.

Gittell, R. and Vidal, A., 1998, *Community Organizing: Building Social Capital as a Development Strategy*, Thousand Oaks: Sage Publications.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A. and Soutter, C. L., 2000, “Measuring Trust”, *The Quarterly Journal of Economics*, Vol. 115(3): 811-846.

Glaeser, E. L., 2001, “The Formation of Social Capital”, in: Helliwell, J.F. (Ed.), The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being. Proceedings of an OECD/HRDC conference, Quebec, March 19–21, 2000. HDRC, Ottawa.

Goodman, E., McEwen, B. S., Huang, B., Dolan, L. M. and Adler, N. E., 2005, “Social Inequalities in Biomarkers of Cardiovascular Risk in Adolescence”, *Psychosomatic Medicine*, Vol. 67: 9-15.

Green, S., 2003, “Performing China’s Economy: A Rough Guide”, the China project, Royal Institute of International Affairs.

Greene, W., 2003, *Econometric Analysis*, 5th ed. Prentice Hall: New Jersey.

Greene, W., 2004, “Fixed Effects and Bias due to the Incidental Parameters Problem in the Tobit Model”, *Econometric Reviews*, Vol. 23: 125-147.

Grogan, C. M., 1995, “Urban Economic Reform and Access to Health Care Coverage in the People's Republic of China”, *Social Science and Medicine*, Vol. 41(8): 1073-1084.

Groot, W. and Maassen van de Brink, H., 2006, “What does Education Do to Our Health?”, in *Measuring the Effects of Education on Health and Civic Engagement*, Proceedings of the Copenhagen Symposium, Centre for Educational Research and Innovation, OECD.

Groot, W. and Maassen van de Brink, H., 2007, “The Health Effects of Education”, *Economics of Education Review*, Vol. 26(2): 186-200.

Groot, W. and Oosterbeek, H., 1994, “Earnings Effects of Different Components of Schooling; Human Capital versus Screening”, *Review of Economics and Statistics*, Vol. 76 (2): 317-321.

Grossman, M., 1972, “On the Concept of Health Capital and the Demand for Health”, *Journal of Political Economy*, Vol. 80: 223-255.

Grossman, M., 1975, “The Correlation between Health and Schooling”, *Household Production and Consumption*, edited by Nestor E. Terleckyj, New York: Columbia University Press, 147-211.

Grossman, M., 2000, “The Human Capital Model of the Demand for Health”. In Culyer, A. J. & Newhouse, J. P. (Eds.), *Handbook of Health Economics*, 1A, 347-408. Amsterdam: Elsevier Science.

Grossman, M., 2006, “Education and Nonmarket Outcomes”, in *Handbook of The Economics of Education*, *Vol. 1*, edited by Eric Hanushek and Finis Welch, North-Holland.

Grossman, M. and Kaestner, R., 1997, “Effects of Education on Health’, in: Behrman, J. & Stacey, N. (eds.), *The Social Benefits of Education*, University of Michigan Press, Ann Arbor, 69-123.

Guo, B. G., 2003, “Transforming China’s Urban Health-Care System”, *Asian Survey*, Vol. 43(2): 385-403.

Guo, S. J., 2003, “The Ownership Reform in China: What Direction and How Far?”, *Journal of Contemporary China*, Vol. 12(36): 553-573.

Habermalz, S., 2003, “Job Matching and the Returns to Educational Signals”, IZA Discussion Paper, No. 726.

Halpern, D., 2005, *Social Capital*, Polity Press, Cambridge, UK.

Han, T., 1998, “China: A shared Poverty to Uneven Wealth”, research project, ECON 270, *Chinese Economy*.

Han, C. and Schmidt, P., 2001, “The Asymptotic Distribution of the Instrumental Variable Estimators When the Instruments are not Correlated with the Regressors”, *Economics Letters*, Vol. 74(1): 61-66.

Hanushek, E. A. and Welch, F., 2006, *Handbook of the Economics of Education*, Vol. 1, Amsterdam; London: North-Holland.

Harmon, C., Oosterbeek, H. and Walker, I., 2003, “The Returns to Education: Microeconomics”, *Journal of Economic Survey*, Vol. 17(2): 115-156.

Hausman, J. A., 1978, “Specification Tests in Econometrics”, *Econometrica*, Vol. 46(6): 1251-1271.

Hay, D. I., 1988, “Socioeconomic Status and Health Status: A study of Males in the Canada Health Survey”, *Social Science and Medicine*, Vol. 27: 1317-1325.

Hayo, B. and Seifert W., 2003, “Subjective Economic Well-being in Eastern Europe”, *Journal of Economic Psychology*, Vol. 24(3): 329-348.

Helliwell, J. F., 2001, “Social Capital, the Economy and Well-Being”, in Keith Banting, Andrew Sharpe and France St-Hilaire, eds., *the Review of Economic Performance and Social Progress*, Vol. 1: 43-60.

Helliwell, J. F., 2003, “How’s Life? Combining Individual And National Variables To Explain Subjective Well-Being”, *Economic Modelling*, Vol. 20(2): 331-360.

Helliwell, J. F., 2006, “Well-being, Social Capital and Public Policy: What’s New?”, *The Economic Journal*, Vol. 116(510): C34-C45(1).

Hemstrom, O., 2005, “Health Inequalities by Wage Income in Sweden: The role of Work Environment”, *Social Science and Medicine*, Vol. 61(3): 637-647.

Holtmann, A. G., 1972, “On-The-Job Training, Obsolescence, Options, and Retraining”, *Southern Economic Journal*, Vol. 38(3): 414-417.

Hough, J., 1994, “Educational Cost-benefit Analysis”, *Education Economics*, Vol. 2(2): 93-128.

Huang, JR-T., Kuo, C. C. and Kao, A. P., 2003, “The Inequality of Regional Economic Development in China between 1991 and 2001”, *Journal of Chinese Economic and Business Studies*, Vol. 1(3): 273-285.

Hudson, J., 2006, “Institutional Trust and Subjective Well-being across the EU”, *Kyklos*, Vol. 59(1): 43-62.

Idler, E. and Benysmini, Y., 1997, “Self-reported Health and Mortality: A Review of Twenty-seven Community Studies”, *Journal of Health and Social Behaviour*, Vol. 38 (1): 21-37.

Idler, E. and Kasl, S., 1995, “Self-rating of Health: Do They also Predict Change in Functional Ability?”, *Journal of Gerontology: Social Sciences*, Vol. 508: S344-S353.

Imbens, G. W. and Rosenbaum, P. R., 2005, "Robust, Accurate Confidence Intervals with A Weak Instrument: Quarter of Birth and Education", *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 168(1): 109-126.

James, H., 1904, *The Golden Bowl*, New York: C. Scribner's Sons.

Janzen, B. L. and Muhajarine, N., 2003, "Social Role Occupancy, Gender, Income Adequacy, Life stage and Health: A Longitudinal Study of Employed Canadian Men and Women", *Social Science and Medicine*, Vol. 57(8): 1491-1503.

Jeffery, R. W. and French, S. A., 1996, "Socioeconomic Status and Weight Control Practices among 20-to-45-year-old Women", *American Journal of Public Health*, Vol. 86(7): 1005-1010.

Jones, D., Li, C. and Owen, A. L., 2003, "Growth and Regional Inequality in China During the Reform Era", William Davidson Working Paper, No. 561, the William Davidson Institute.

Jones, A. M. and Schurer, S., 2007, "How Does Heterogeneity Shape the Socioeconomic Gradient in Health Satisfaction?", Ruhr Economic Paper, No. 8.

Kahneman, D., Diener, E. and Schwarz, N. (eds), 1999, *Well-being: the Foundations of Hedonic Psychology*. New York: Russell Sage Foundation.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A. and Jaffe, M. W., 1963, "Studies of Illness in the Aged: The Index of ADL, A Standardized Measure of Biological and Psychosocial Function", *Journal of the American Medical Association*, Vol. 185: 914-919.

Kidd, M. P. and Meng X., 2001, “The Chinese State Enterprise Sector: Labour Market Reform and the Impact on Male-female Wage Structure”, *Asian Economic Journal*, Vol. 15: 405-423.

Knight, J. and Gunatilaka, R., 2007, “Great Expectations? The Subjective Well-Being of Rural-Urban Migrants in China”, Discussion Paper series, ISSN 1471-0498.

Knight, J. and Shi, L., 1996, “Educational Attainment and the Rural-Urban Divide in China”, *Oxford Bulletin of Economics and Statistics*, Vol. 58(1): 83-117.

Knight, J. and Song L., 2003, “Increasing Wage Inequality in China: Extent, Elements and Evaluation”, *The Economics of Transition*, Vol. 11(4): 597-619.

Kroch, E. A. and Sjoblom, K., 1994, “Schooling as Human Capital or A Signal: Some Evidence”, *The Journal of Human Resources*, Vol. 29(1): 156-180.

Krueger, A. B. and Pischke, J. S., 1995, “A Comparative Analysis of East and West German Labour Markets: Before and After Unification,” in *Differences and Changes in Wage Structure*, Freeman, R. and Katz, L. (eds.), Chicago: University of Chicago Press.

Lardy, N. R., 2002, *Integrating China into the Global Economy*, Brookings Institution Press.

Laroche, M., Merette, M. and Ruggeri, G. C., 1999, “On the Concept and Dimensions of Human Capital in a Knowledge-Based Economy”, *Canadian Public Policy*, Vol. 25 (1): 87-100.

Lavie, C. J., Kuruvanka, T., Milani, R. V., Prasad, A. and Ventura, H. O., 2004, "Exercise Capacity in Adult African-Americans Referred for Exercise Stress Testing: Is Fitness Affected by Race?", *Chest*, Vol. 126: 1962-1968.

Lawton, M. P. and Brody, E., 1969, "Assessment of Older People: Self-maintaining and Instrumental Activities of Daily Living", *The Gerontologist*, Vol. 9: 179-186.

Layard, R., 2005, *Happiness-Lessons from a New Science*. London: Penguin.

Layard, R. and Psacharopoulos, G., 1974, "The Screening Hypothesis and the Returns to Education", *Journal of Political Economy*, Vol. 82(5): 985-998.

Li, H., 2003, "Economic Reform and Returns to Education in China", *Economics of Education Review*, Vol. 22: 317-328.

Li, H. B., Liu, P. W., Ma, N. and Zhang, J. S., 2005, "Does Education Pay in Urban China? Estimating Returns to Education Using Twins", paper provided by Chinese University of Hong Kong, Department of Economics in its series discussion paper, No. 00013.

Li, Q., de Brauw, A., Rozelle, S. and Zhang, L. X., 2005, "Labour Market Emergence and Returns to Education in Rural China", *Review of Agricultural Economics*, Vol. 27(3): 418-424.

Liang, J., McCarthy, J. F., Jain, A., Krause, N., Bennett, J. M. and Gu, S. Z., 2000, "Socioeconomic Gradient in Old Age Mortality in Wuhan, China", *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, Vol. 55: S222-S233.

Lin, J. Y., 2004, “Lessons of China’s Transition from a Planned Economy to a Market Economy”, WSPiZ and TIGER, Distinguished Lectures Series No. 16, Warsaw.

Lin, J. Y., Cai, F. and Li, Z., 1996, “The Lessons of China’s Transition to a Market Economy”, *Cato Journal*, Vol. 16(2): 201-31.

Liu, Z. Q., 1998, “Earnings, Education and Economic Reforms in Urban China”, *Economic Development and Cultural Change*, Vol. 46(4): 697-725.

Lu, M. and Chen, Z., 2006, “Urbanization, Urban-Biased Policies, and Urban-Rural Inequality in China, 1987-2001”, *Chinese Economy*, Vol. 39(3): 42-63.

Lu, L. and Lin, Y. Y., 1998, “Family Roles and Happiness in Adulthood”, *Personality and Individual Differences*, Vol. 25: 195-207.

Mailath, G. J., 1987, “Incentive Compatibility in Signaling Games with a Continuum of Types”, *Economerica*, Vol. 55(6): 1349-1366.

Maluccio, J., 1998, “Endogeneity of Schooling in the Wage Function: Evidence from the Rural Philippines”, FCND Discussion Paper, No. 54.

Matthews, S., Hertzman, C., Ostry, A. and Power, C., 1998, “Gender, Work Roles and Psychosocial Work Characteristics as Determinants of Health”, *Social Science and Medicine*, Vol. 46(11): 1417-1424.

Matthews S., Manor, O. and Power, C., 1999, “Social Inequalities in Health: Are There Gender Differences?’, *Social Science and Medicine*, Vol. 48(1): 49-60.

Maurer-Fazio, M., 1999, “Earnings and Education in China’s Transition to a Market Economy Survey Evidence from 1989 and 1992”, *China Economic Review*, Vol. 10(1): 17-40.

McBride, M., 2001, “Relative-income Effects on Subjective Wellbeing in the Cross-section”, *Journal of Economic Behaviour and Organization*, Vol. 45: 251-278.

Meer, J., Miller, D. and Rosen, H., 2003, “Exploring the Health-wealth Nexus”, *Journal of Health Economics*, Vol. 22 (5): 713-730.

Mehnert, T., Krauss, H. H., Nadler, R. and Boyd, M., 1990, “Correlates of Life Satisfaction in Those with Disabling Conditions”, *Rehabilitation Psychology*, Vol. 35: 3-17.

Mehta, A. and Villarreal, H. J., 2004, “What Causes the Sheepskin Effect? An Expanded Mincerian Framework Applied to Mexico”, *University of Wisconsin-Madison Working Paper*.

Meng, X., 2000, *Labour Market Reform in China*, Cambridge University Press, Cambridge, England.

Meng, X. and Zhang, J. S., 2001, “The Two-Tier Labour Market in Urban China”, *Journal of Comparative Economics*, Vol. 29: 485-504.

Mincer, J., 1974, *Schooling, Experience and Earnings*, New York: National Bureau of Economic Research.

Mokdad, A. H., Serdula, M. K., Dietz, W. H., Marks J. S. and Koplan J. P., 1999, "The Spread of the Obesity Epidemic in the United States, 1991–1998", *Journal of the American Medical Association*, Vol. 282(16): 1519-1522.

Monden, C. W., van Lenthe, F., de Graaf, N. D. and Kraaykamp, G., 2003, "Partner's and Own Education: Does Who You Live with Matter for Self-assessed Health, Smoking and Excessive Alcohol Consumption?", *Social Science and Medicine*, Vol. 57(10): 1901-1912.

Montgomery, S. M., Cook, D. G., Bartley, M. J. and Wadsworth, M. E. J., 1998, "Unemployment, Cigarette Smoking, Alcohol Consumption and Body Weight in Young British Men", *European Journal of Public Health*, Vol. 8(1): 21-27.

Moon, O. R., Kim, N. S., Jang, S. M., Yoon, T. H. and Kim, S. O., 2002, "The Relationship between Body Mass Index and the Prevalence of Obesity-related Diseases based on the 1995 National Health Interview Survey in Korea", *Obesity Reviews*, Vol. 3(3): 191-196.

Mulatu, M. S. and Schooler, C., 2002, "Causal Connections between Socio-economic Status and Health: Reciprocal Effects and Mediating Mechanisms", *Journal of Health and Social Behaviour*, Vol. 43(1): 22-41.

Mundlak, Y., 1978, "On the Pooling of Time Series and Cross-section Data", *Econometrica*, Vol. 46(1): 69-85.

Myers, D., 1999, “Close Relationships and Quality of Life”, in D. Kahneman, E. Diener and E. Schwartz (eds.), *Well-being: The Foundations of Hedonic Psychology*, Russell Sage Foundation, New York.

Namazie, C. and Sanfey, P., 1999, “Happiness in Transition: the Case of Kyrgyzstan”, Discussion Paper, DARP 40.

Neuman, S. and Weiss, A., 1995, “On the Effects of Schooling Vintage on Experience-Earnings Profiles: Theory and Evidence”, *European Economic Review*, Vol. 39(5): 943-955.

Neumark, D., 1999, “Biases in Twin Estimates of the Return to Schooling”, *Economics of Education Review*, Vol. 18(2): 143-148.

Nie, N. H., Junn, J. and Stehlik-Barry, K., 1996, *Education and Democratic Citizenship in America*, Chicago: University of Chicago Press.

Organization for Economic Cooperation and Development (OECD), 2001, *The Well-Being of Nations: The Role of Human and Social Capital*. Paris: OECD Centre for Educational Research and Innovation.

Okushima, S. and Uchimura, H., 2005, “Economic Reforms and Income Inequality in Urban China”, Institute of Developing Economies Discussion Paper, No. 25.

Opper, S., 2001, “Dual-track Ownership Reforms: Lessons from Structural Change in China, 1978-1997”, *Post-Communist Economies*, Vol. 13(2): 206-227.

Oswald, A. J., 1997, “Happiness and Economic Performance”, *Economic Journal*, Vol. 107(445): 1815-1831.

Park, J., 1999, “Estimation of Sheepskin Effects Using the Old and New Measures of Educational Attainment in the Current Population Survey”, *Economics Letters*, Vol. 62(3): 237-240.

Persell, C. H., Green, A. and Gurevich, L., 2001, “Civil Society, Economic Distress, and Social Tolerance”, *Sociological Forum*, Vol. 16(2): 203-230.

Peterson, C. and Stunkard, A. J., 1989, “Personal Control and Health Promotion”, *Social Science in Medicine*, Vol. 28: 819-828.

Prentice, A. M. and Jebb, S. A., 2001, “Beyond body mass index”, *Obesity Reviews*, Vol. 2(3): 141-147.

Psacharopoulos, G., 1979, “On the Weak versus the Strong Version of the Screening Hypothesis”, *Economics Letters*, Vol. 4: 181-185.

Psacharopoulos, G. 1985, “Returns to Education: A Further International Update and Implications”, *Journal of Human Resources*, Vol. 20(4): 583-597.

Psacharopoulos, G. 1994, “Returns to investment in Education: A Global Update”, *World Development*, Vol. 22(9): 1325-1343.

Psacharopoulos, G. and Patrinos, H. A., 2002, "Returns to Investment in Education: A Further Update", World Bank Policy Research Working Paper, No. 2881, Washington, D.C.: World Bank.

Putnam, R. D., Leonardi, R. and Nanetti, R. Y., 1995, *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Putnam, R. D., 1995, "Bowling Alone: America's Declining Social Capital", *Journal of Democracy*, Vol. 6(1): 65-78.

Putnam, R. D., 2000, *Bowling Alone: The Collapse and Revival of American Community*. Simon & Schuster, New York.

Putnam, R. D., 2001. "Social Capital: Measurement and Consequences", in: Helliwell, J.F. (Ed.), The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being. Proceedings of an OECD/HRDC conference, Quebec, March 19–21, 2000. HDRC, Ottawa.

Qian, S. M., 1997, "The Process and Periods of the Wage System Transition in the State-Owned Enterprises", *The Economists* (*Jingji Xuejia*), in Chinese, Vol. 2: 32-37.

Qian, X. L. and Smyth, R., 2005, "Measuring Regional Inequality of Education in China: Widening Coast-Inland Gap or Widening Rural-Urban Gap?", ABERU Discussion Paper, Monash University.

Ramirez, J. V., 2002, "Age and Schooling Vintage Effects on Earnings Profiles in Switzerland", in: De Grip, A., J. van Loo and K. Mayhew (eds.), *The Economics of Skills*

Obsolescence, Research in Labour Economics, Vol. 21: 83-99, Amsterdam /Boston: JAI Press.

Rawski, T., 2005, "Recent Developments in China's Labour Economy", Working Paper, No. 280, Department of Economics, University of Pittsburgh.

Redding, S. and Venables, A. J., 2004, "Economic Geography and International Inequality", *Journal of International Economics*, Vol. 62: 53-82.

Reddy, B. N., 1998, "Body Mass Index and its Association with Socioeconomic and Behavioural Variables among Socioeconomically Heterogeneous Population of Andhra Pradesh, India", *Human Biology*, Vol. 70(5): 901-917.

Rice, P. and Venables, A. J., 2003, "Equilibrium Regional Disparities: Theory and British Evidence", *Regional Studies*, Vol. 37(6/7): 675-686.

Riley, J., 1975, "Competitive Signaling", *Journal of Economic Theory*, Vol. 10: 174-86.

Riley, J., 1979, "Testing the Educational Screening Hypothesis", *Journal of Political Economy*, Vol. 87(5): Part 2, S227-252.

Rong, X. L. and Shi, T., 2001, "Inequality in Chinese Education", *Journal of Contemporary China*, Vol. 10(26): 107-124.

Rosen, S., 1975, "Measuring the Obsolescence of Knowledge", in Juster, F. T. (ed.), *Education, Income and Human Behaviour*, 199-232, New York: McGraw-Hill.

Roy, J. R. and Thill, J. C., 2004, “Spatial Interaction Modelling”, *Papers in Regional Science*, Vol. 83(1): 339-361.

Sander, W., 1995, “Schooling and Smoking”, *Economics of Education Review*, Vol. 14(1): 23-33.

Sanfey, P. and Teksoz, U., 2006, “Does Transition Make You Happy?”, Working Paper series, No. 91, EBRD.

Schultz, T. W., 1961, “Investment in Human Capital”, *The American Economic Review*, Vol. 51(1): 1-17.

Scitovsky, T., 1976, *The Joyless Economy: An Inquiry Into Human Satisfaction and Dissatisfaction*. Oxford U.P., Oxford.

Shan, C., 1991, “Implementation of a New Within-firm Distribution System”, *The Reforms in Labour, Wage, and Social Security System*, in Chinese, China’s Labour Press, Beijing.

Shek, T. L., Chan, Y. K. and Lee, P. S. N., 2005, “Quality of Life in the Global Context: A Chinese Response”, *Social Indicators Research*, Vol. 71(1-3).

Shen, S. M. and Choy, S. T. B., 2005, “The Pre- and Post-1997 Well-Being of Hong Kong Residents”, *Social Indicators Research*, Vol. 71(1-3).

Shi, L. Y., 1993, “Health care in China: a rural-urban comparison after the socioeconomic reforms”, *Bulletin of the World Health Organization*, Vol. 71(6): 723-736.

Shields, M. A. and Price, S. W., 2005, “Exploring the Economic and Social Determinants of Psychological Well-being and Perceived Social Support in England”, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 168(3): 513-537.

Sicular, T., Yue, X. M., Li., S. and Gustafsson, B., 2005, “The Urban-Rural Gap and Income Inequality in China”, paper presented for UNU-WIDER Project Meeting “*Inequality and Poverty in China”*, 26-27 August, Helsinki, Finland.

Skinner, E. A., 1996, “A Guide to Constructs of Control”, *Journal of Personal and Social Psychology*, Vol. 71: 546-570.

Skirbekk, V., 2002, “Variations in Productivity over the Life Span: A Review and Some Implications”, revised versions of IIASA Interim Report IR-02-061, International Institute for Applied Systems Analysis.

Smith, A., 1776, *The Wealth of Nations*, edited by Edwin Cannan, 1994 Modern Library Edition, Random House Inc.

Smith, J. P., 1999, “Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status”, *The Journal of Economic Perspectives*, Vol. 13(2): 145-166.

Smith, J. P., 2004, “Unravelling the SES-Health Connection”, IFS Working Papers, W04/02, Institute for Fiscal Studies.

Sobal, J. and Stunkard, A. J., 1989, “Socioeconomic Status and Obesity: A Review of the Literature”, *Psychological Bulletin*, Vol. 105(2): 260-275.

Spasojevic, J., 2003, "Effects of Education on Adult Health in Sweden: Results from a Natural Experiment", Mimeo, The Graduate School for Public Affairs and Administration, Metropolitan College of New York.

Spence, A. M., 1973, "Job Market Signaling", *Quarterly Journal of Economics*, Vol. 87: 355-74.

Spence, A. M., 1974, *Market Signaling*, Cambridge: Harvard University Press.

Stiglitz, J. E., 1975, "The Theory of 'Screening', Education, and the Distribution of Income", *American Economic Review*, Vol. 65: 283-300.

Stock, J. H., Wright, J. H. and Yogo, M., 2002, "A survey of Weak Instruments and Weak Identification in Generalized Method of Moments", *Journal of Business and Economic Statistics*, Vol. 20(4): 519-529.

Stutzer, A., 2004, "The Role of Income Aspirations in Individual Happiness", *Journal of Economic Behaviours and Organization*, Vol. 54: 89-109.

Su, T. T., 1973, "Skill Obsolescence and Training: Comment", *Southern Economic Journal*, Vol. 39(4): 645-648.

Syme, S. L., 1989, "Control and Health: A Personal Perspective", *Stress, Personal Control and Health*, eds. A. Steptoe and A. Appels, 3-18, John Wiley and Sons, Brussels.

Tsang, M., 2000, "Education and National Development in China since 1949: Oscillating Policies and Enduring Dilemmas", *China Review 2000*: 579-618.

Tobin, J., 1958, “Estimation of Relationships for Limited Dependent Variables”, *Econometrica*, Vol. 26: 24-36.

Uchimura, H., 2005, “Impact of Changes in Social Institutions on Income Inequality in China”, OECD Development Centre Working Papers, No. 243, OECD Development Centre.

Van Doorslaer, E., *et al.*, 2000, “Equity in the Delivery of Health Care in Europe and the US”, *Journal of Health Economics*, Vol. 19: 553-583.

Van Doorslaer, E. and Koolman, X., 2002, “Explaining the Differences in Income-related Health Inequalities across European Countries”, *ECuity II Project*, Working Paper, No. 6.

Van Imhoff, E., 1988, “Age Structure, Education, and the Transmission of Technical Change”, *Journal of Population Economics*, Vol. 1(3): 167-181.

Verhoeven, W. J., Jansen W. and Dessens, J., 2005, **“**Income Attainment during Transformation Processes: A Meta-Analysis of the Market Transition Theory”, *European Sociological Review*, Vol. 21(3): 201-226.

Wang, Y., Mi, J., Shan, X., Wang, Q. J. and Ge. K., 2007, “Is China Facing an Obesity Epidemic and the Consequences? The Trends in Obesity and Chronic Disease in China”, *International Journal of Obesity*, Vol. 31: 177-188.

Wardle, J., Robb, K. A., Johnson, F., Griffith, J., Brunner, E., Power C. and Tovee, M., 2004, “Socioeconomic Variation in Attitudes to Eating and Weight in Female Adolescents”, *Health Psychology*, Vol. 23(3): 275-282.

Wei, Y. Q., 2003/2, “Foreign Direct Investment in China”, Working Paper, Lancaster University Management School.

Weiss, Y., 1971, “Ability and the Investment in Schooling: A Theoretical Note on J. Mincer’s ‘Distribution of Labor Income’”, *Journal of Economic Literature*, Vol. 9(2): 459-461.

Weiss, A., 1995, “Human Capital vs. Signalling Explanations of Wages”, *Journal of Economic Perspective*, Vol. 9(4): 133-154.

WHO Expert Consultation, 2004, “Appropriate Body-mass Index for Asian Populations and its Implications for Policy and Intervention Strategies”, *Lancet*, Vol. 363: 157-163.

Wildman, R. P., Gu, D. F., Reynolds, K., Duan, X. F. and He, J., 2004, “Appropriate Body Mass Index and Waist Circumference Cutoffs for Categorization of Overweight and Central Adiposity among Chinese Adults”, *American Journal of Clinical Nutrition*, Vol. 80(5): 1129-1136.

Winkleby, M. A., Jatulis, D. E., Frank, E. and Fortmann, S. P., 1992, “Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease”, *American Journal of Public Health*, Vol. 82(6): 816-820.

Woolcock, M., 2001, “The Place of Social Capital in Understanding Social and Economic Outcomes”, *Isuma: Canadian Journal of Policy Research*, Vol. 2(1): 1-17.

Wooldridge, J., 2002, *Econometric Analysis of Cross Section and Panel Data*, the MIT Press.

Wu, R. T., 2003, *Medical Insurance Systems: An International Comparison,* in Chinese, Huaxue Gongye Press, Beijing.

Wu X. G. and Xie, Y., 2002, “Does the Market Pay Off? Earning Inequality and Returns to Education in Urban China,” William Davidson Working Paper, No. 454.

Yi, Z., Vaupel, J. W., Xiao, Z. Y., Zhang, C. Y. and Liu, Y. Z., 2002, “Sociodemographic and Health Profiles of the Oldest Old in China’, *Population and Development Review*, Vol. 28(2): 251-273.

Yu, X. and Hannum, E., 1996, “Regional Variation in Earnings Inequality in Reform-Era Urban China”, *America Journal of Sociology,* Vol. 101(4): 950-992.

Yueh, L., 2004, “Wage Reforms in China during the 1990s”. *Asian Economic Journal*, Vol. 18(2): 149-164.

Zhang L. W., 2005, “Prediction of Chinese Life Satisfaction: Contribution of Collective Self-esteem”, *International Journal of Psychology*, Vol. 40(3): 189-200.

Zhang, X. B. and Kanbur, R., 2005, “Spatial Inequality in Education and Health Care in China”, *China Economic Review*, Vol. 16(2): 189-204.

Zhang, J. S., Zhao, Y. H., Park, A. and Song X. Q., 2005, “Economic Returns to Schooling in Urban China, 1988-2001”, *Journal of Comparative Economics,* Vol. 33(4): 730-752.

Zhao, Y. H., 2002, “Earnings Differentials between State and Non-state Enterprises in Urban China”, *Pacific Economic Review*, Vol. 7(1): 181-97.

Zhao, Z., 2005, “Health Determinants in Urban China”, Discussion Paper series, IZA DP No. 1835.

Zhao, Z. W., 2006, “Income Inequality, Unequal Health Care Access, and Mortality in China”, *Population and Development Review*, Vol. 32(3): 461-483.

Zhao, W. and Zhou, X. G., 2002, “Institutional Transition and Returns to Education in Urban China: An Empirical Assessment”, *Research in Social Stratification and Mobility*, Vol. 19: 337-373.

Zhu, S. Y. and Zhang, S. K., 1987, “Development of Primary Health Care in Chinese Rural Areas under Socialist Economy Reform”, in Chinese, *China Rural Health Service Administration*, Vol. 22.

Zimmer, Z., Chayovan, N., Lin, H. S. and Natividad. J. N., 2004, “How Indicators of Socioeconomic Status Relate to Physical Functioning of Older Adults in Three Asian Societies”, *Research on Aging*, Vol. 26(2): 224-258.

Zimmer, Z. and Kwong, J., 2004, “Socioeconomic Status and Health among Older Adults in Rural and Urban China”, *Journal of Aging and Health*, Vol. 16 (1): 44-70.