UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

School of Social Sciences, Department of Economics

Three Essays on Migration in China

by

Hao Xu

A thesis submitted for the degree of Doctor of Philosophy

September 2016

To my loved ones.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES
School of Social Sciences, Department of Economics

Doctor of Philosophy

Three Essays on Migration in China

by Hao Xu

This thesis focuses on the impacts of rural-urban migration on economic development in China. It includes three papers studying the impact of migration and return migration in China. The first chapter of the thesis introduces the background, reviews studies in the field and shows the contribution of the papers and potential policy implications. The second chapter is about time use and labour supply of the left behind children and spouse of migrants. The third chapter is on the impact of return migration on consumption behaviour, the fourth chapter is about the impact of mass media and migration on the diet knowledge of rural residents, while the last chapter concludes the thesis.

The rural to urban migration in China is the largest population movement in the world history Zhao (1999). Massive migration emerged at the beginning of the 1980s with the reform and open policy. According to Meng (2012), the number of rural to urban migrants is estimated to increase by more than 100 million. Migrants in China always leave their spouses and children in villages, and then return to the villages after working in the urban areas temporarily. This temporary deployment is due to the household registration system (which segregates rural and urban areas). Over the last few decades, during the rapid rural development and urbanisation in China, there have been many changes in a variety of parameters in Chinese villages. The study investigates questions like, how does the traditionally closed, but developing villages transform due to the inevitable urbanisation process; how does the rural families that have been left behind, adjust with the household time utilisation, when facing migration; and also how do the rural residents change their consumption behaviours, social norms and knowledge.

Motivated by the changes mentioned above, this thesis aims to provide answers to the following research questions - How does the time utilisation of left behind family members affected by the migration of the male household head? How does migration impacts the behaviour and knowledge of individuals? While the first paper discusses the effect of migration on the time utilisation of the left behind children and women in rural China, the second paper evaluates the impact of migration and return migration on the smoking and drinking behaviour of the rural males and the third paper investigates its effect on the diet knowledge of rural individuals.

There are only a few studies about the influence of rural-urban migration on the individuals left behind, particularly regarding their time use and labour supply. The migration of a household member might lead to an income effect that could reduce the labour supply of the left behind and/or to a substitution effect, which would work contrarily. Thus, the comprehensive effect of migration on the left behind is an empirical question. The first paper utilises the panel data from the China Health and Nutrition Survey (CHNS), and controls for the endogeneity of migration, using the instrumental variable approach. The findings indicate that the time allocation of the left behind spouse and children is not statistically altered from the same members of non-migrant households. This suggests that the income effect is offset by the substitution effect and the result proves to be resilient to various robust checks.

As smoking and drinking are common behaviours in China, the second paper investigates how the return migration impacts the smoking and drinking behaviour of the migrants. By controlling for the self-selection of migration, and comparing the smoking behaviour before and after the migration, the second paper identifies that, generally, return migrants are more likely to smoke and drink than non-migrants and in higher frequency. This positive and significant effect is driven by males. After controlling for the effect of income, the impact of return migration on smoking and drinking is found to be consistent. The changes in males' smoking and drinking pattern, before and after the migration suggests that it is not the income effect, but the migration experience that makes the male returnees more inclined towards smoking and drinking. The results present suggestive evidence that migrants adapt to the social norm in urban areas and

bring these behaviours back to villages, upon return.

The third paper studies the transfer of diet knowledge in Chinese villages, via migration and mass media as the two channels of information flow. It probes whether exposure to migration and/or mass media impacts the dietary knowledge of the rural residents. Panel fixed effects analysis of unique data containing information on diet knowledge and mass media indicates that individual television exposure and migration density in the village has a great and significant impact on the diet knowledge of the rural residents. The study demonstrates significant spillover effect of migration density and strong peer effect. The variation of aggregate mass media exposure in villages over the years is used as an instrumental variable to deal with the endogeneity of individual mass media in the model. Furthermore, the results are consistent when media exposure is measured as Internet exposure.

This thesis contributes to the growing, but the limited literature on empirical development studies, which are focusing on rural to urban migration in the developing countries. The contribution of each section in the study is discussed separately in the papers. In the recent years, it has been observed that migration has been a hot research field in the development studies, both in international migration and internal migration within the developing countries. This thesis comprises of three papers that focus on the internal migration in China. These three sections provide empirical evidence on the rural to urban migration in China and explain the impact of migration on the individuals left behind, the return migrants and on the rural non-migrant residents. This thesis also contributes to the emerging literature regarding the impact of migration on social norms and knowledge. The findings of this thesis offer effective policy implication for the policy makers, in order to maximise the benefits of migration, for better understanding about the rural development in China as well as for poverty reduction and urbanisation.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, Hao Xu, declare that the thesis entitled *Three Essays on Migration in China* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission;

Signed: .....

Date: .....

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my Ph.D. supervisors, Professor Jackie Wahba and Dr. Christian Schluter, for their immense knowledge, suggestions, and insights. I am immensely grateful to have them as my supervisors and their guidance and wisdom will continue to influence my career and life in the future. I would especially like to thank Jackie for being patient, and leading my way through research with her innovative ideas. Her inspiration is the fundamental basis of the two chapters in this thesis - Chapter 3, "Return Migration and Consumption Behaviour", and Chapter 4, "Migration, Mass Media and Diet Knowledge".

I appreciate the help that I received from the Economics department and would like to acknowledge the nice and friendly staff for their assistance. I would also like to extend my heartfelt thanks to Michael Vlassopoulos, Helen Paul, Jian Tong, Corrado Giulietti and others. Thanks to Michael Kearns, Carlos Pineda, Tong Xue, Jana Sadeh, Chuhong Wang, Panos Ginnarakis, Dafni Papoutsaki, Michele Tuccio, Andres Luengo, Elif Kara and others for their constant support, friendship, and concern.

I am also indebted to the Economic and Social Research Council (ESRC) for their

ESRC DTC +3 Studentship and Research Training Support Grant (RTSG) that funded this research from 2012 to 2015.

This research utilizes the data from the China Health and Nutrition Survey (CHNS) and acknowledges the contribution of the National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center (5 R24 HD050924), the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, R24 HD050924, and R01-HD38700) and the Fogarty International Center. A special gratitude for the NIH that offered financial support for the CHNS data collection and examining the files from 1989 to 2011 along with future surveys, and also to the China-Japan Friendship Hospital, Ministry of Health for its support for CHNS 2009.

1

Introduction

1.1 Background

In recent years, it has been observed that research in migration has constituted to be an important topic in development studies. Both internationally and internally in China, the stock and flow of human capital has been an area of focus in economic research. This thesis consists of three papers, which study the novel research questions on the impact of migration and return migration in China. The papers present the research with an emphasis on time use and labour supply of the left behind family members, the impact of return migration on consumption behaviour, mass media, migration and diet knowledge in Chinese villages.

This thesis is comprised of three papers that investigate the novel research questions regarding the impact of migration and return migration in China. The papers examine the concept of time use and labour supply of the left behind family members, the im-

pact of return migration on the consumption behaviour, mass media, migration and diet knowledge in Chinese villages.

Over the last decades, during the rapid rural development and urbanisation in China, there have been many changes in a variety of parameters in Chinese villages. The research attempts to provide answers to questions like, how does the rural left behind family adjust the time allocation when facing migration? How do the rural residents change their consumption behaviours, norms and knowledge? This thesis is motivated by the dramatic changes in Chinese rural villages.

The study is presented in various sections. The introduction section of the thesis, describes the background, reviews literature in the field, introduces the data and methodology and presents the contribution and relevant links of the papers in this thesis. The first section provides the background of the thesis, followed by the second section that reviews the literature on the impact of migration on the left behind rural residents. In the third section, a survey of literature on the impact of return migration and the changes in social norms is discussed. The fourth section describes the data used in this thesis, followed by the fifth section that explains the econometric methods in the empirical studies. Lastly, the sixth part analyses the interrelations among the three papers, their position in the field and states the innovation and contribution of the thesis to the focus of the research.

As per the available statistics, the rural to urban migration in China is the largest population mobility in world history Zhao (1999). Figure 1.1 as cited from the Reserve Bank of Australia[1] depicts the number of rural migrants since the 1990s. The number of rural migrants in 2010 were over 150 million and the increment is more than 100 million, which is consistent with Meng (2012). The concept of rural to urban migration in China is one of the most significant phenomena in the last decades at the transition period with high GDP growth.

The population of China is undergoing a change. A set of figures based on the data from the World Bank presents the big picture of the population, urbanisation and rural

[1] http://www.rba.gov.au/publications/bulletin/2011/sep/4.html

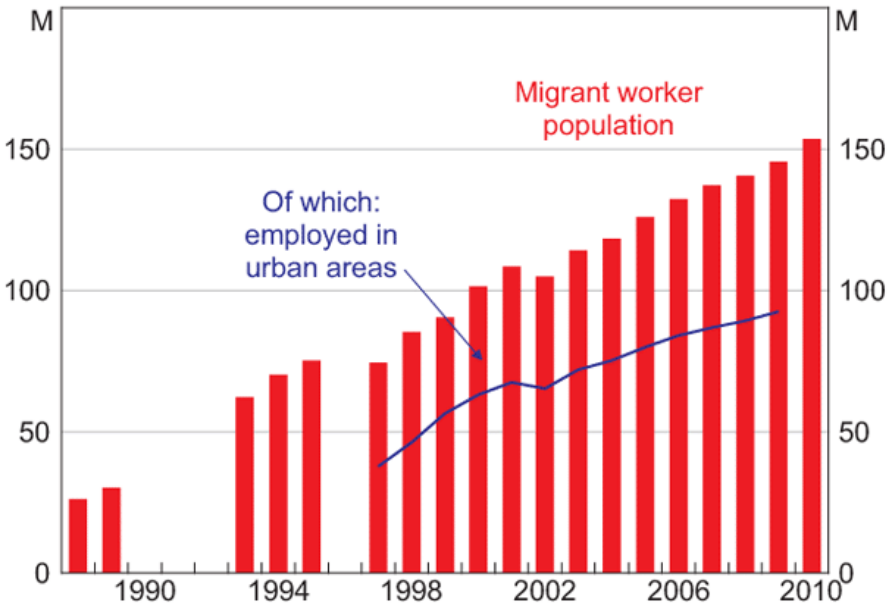


Figure 1.1: China rural to urban migration

development. Figure 1.2 depicts the proportion of the rural and urban population in China from the 1990s to 2014, and figure 1.3 illustrates the population growth rates, accordingly. It was seen that around 2010, the urban population was greater than the rural population. Figure 1.4 presents the decrease of agriculture in GDP and the increase in service. Similarly, figure 1.5 shows the decreasing trend in the population size of young (0-14) generation and predicts a rapid increase in ageing population in the near future.

Figure 1.6 to figure 1.9 are based on the World Bank data and show some development aspects in China over the period. Also, this urbanisation is apparent in Figure 1.6, which shows the rapid growth in mobile phone and Internet subscribers. In addition to migration, chapter four of this thesis would also evaluate the impact of mass media on the rural individuals. Figure 1.7 shows the cluster of population in big cities with more than 1 million residents and the decrease of slum households in urban areas. Figure 1.8 depicts the rural population with access to electricity, non-solid fuel, improved sanitation facilities and improved drinking water source. Also, Figure 1.9 displays the

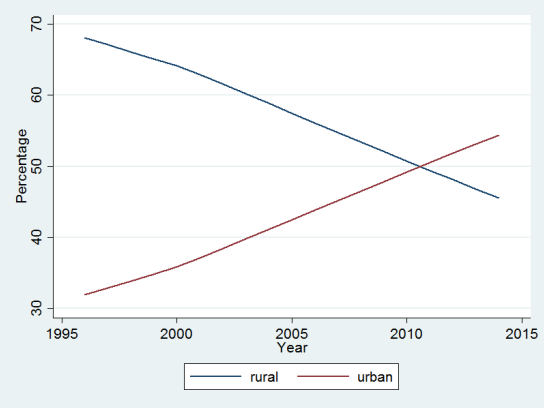


Figure 1.2: Population in urban and rural areas

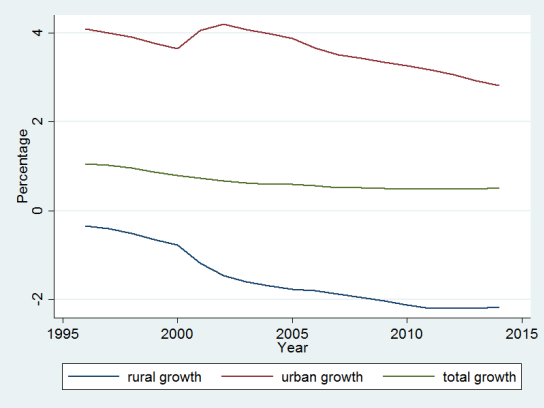


Figure 1.3: Population growth in urban and rural areas

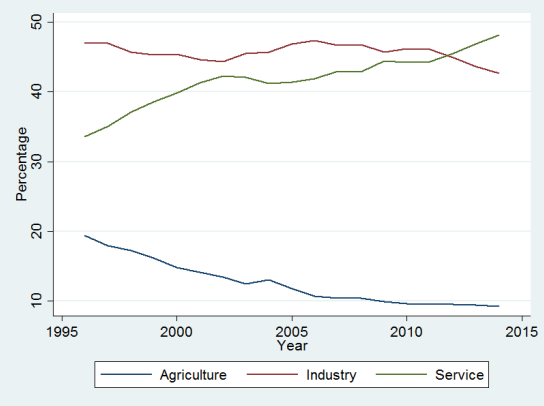


Figure 1.4: Share of value added of GDP by sector

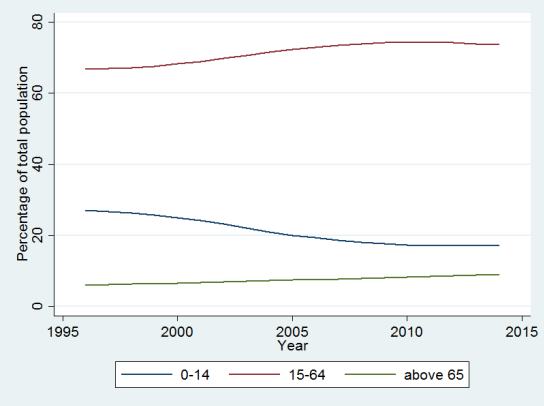


Figure 1.5: Population structure

agriculture value added per agriculture worker in 2005.

Although there is rapid development in both the urban and rural areas in China, there are still many social and economic issues prevalent. For instance, with the segregation of the rural and urban areas, the household registration system (Hukou) restricts rural residents from moving freely and accessing the public services and benefits in urban areas. Migrants in China always leave their spouse and children in villages. Estimated from the Chinese census, 61 million children were left behind in 2010, which amounts

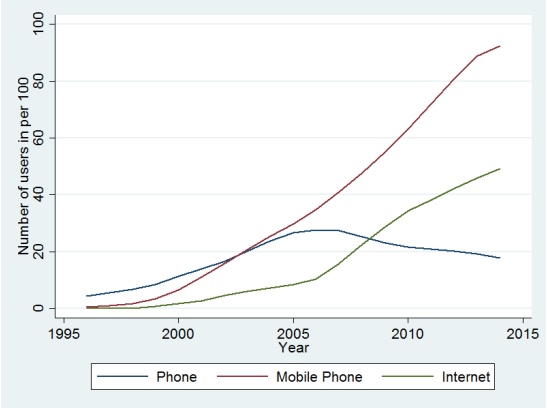


Figure 1.6: Phone and Internet users

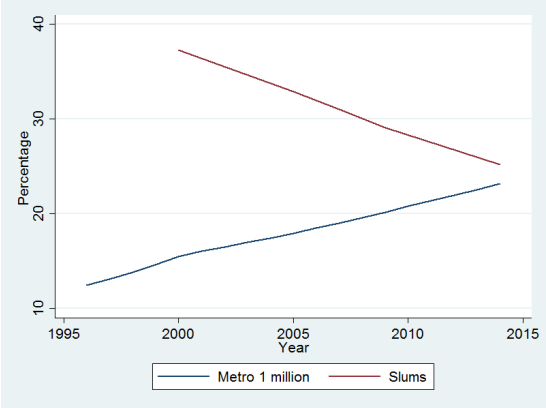


Figure 1.7: Cluster in large cities

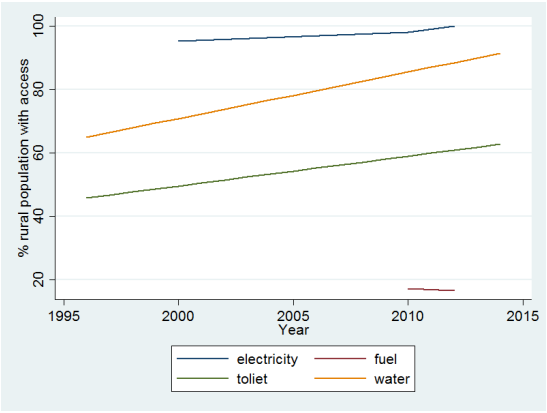


Figure 1.8: Rural development

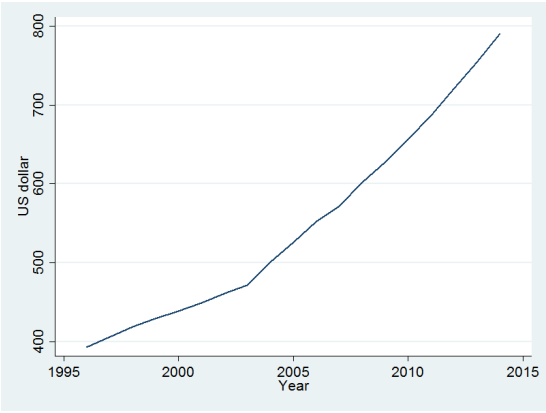


Figure 1.9: Agriculture productivity per farmer

to the entire population of the UK[2]. There were around 50 million left behind women in 2006 (Zhang (2006)). In fact, migrants are compelled to return to the villages after temporarily working in the urban areas, because of this household registration system, which has segregated rural and urban areas in the last decades. However, in the recent years, since 2010, there has been some change in the policy. Chinese authorities has started encouraging the rural residents to settle in local urban areas for numerous reasons.

[2]Chinese official news agency, The people's daily http://acwf.people.com.cn/n/2013/0510/c99013-21437965.html

Rural to urban migration has a strong tendency to cluster in certain sectors and geographical regions. Most rural to urban migration are inter-province movements, which goes from inland provinces to coastal urban areas. The following maps from the Reserve Bank of Australia[3] displays the sending (figure 1.10) and receiving provinces (figure 1.11).

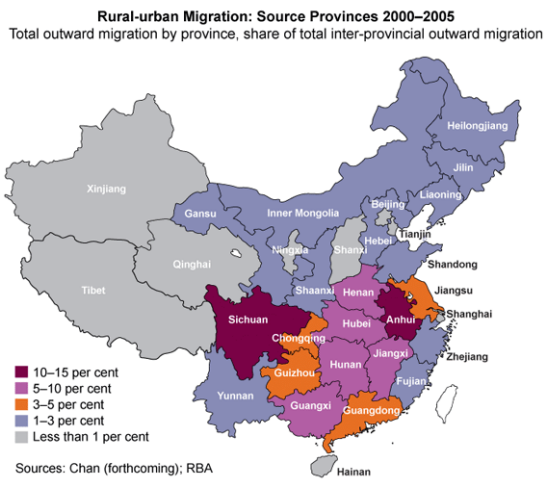


Figure 1.10: Migration sending provinces

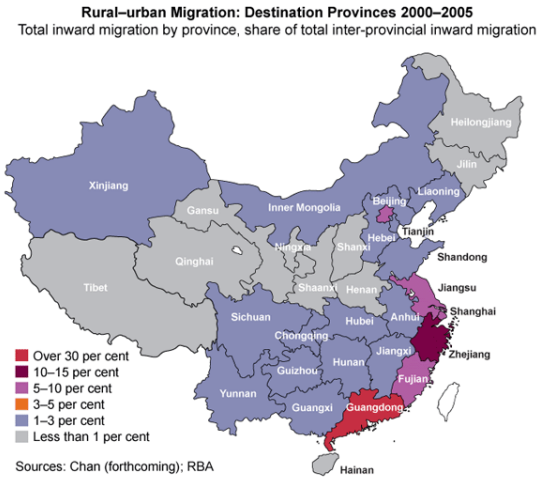


Figure 1.11: Migration destinations

1.2 Survey on Migration and the Left Behind

For a long time, the study of migration has been limited to the impact of migration on the migrants, the destination of relocation, determinants of migration and return migration. Recently, theories about migration and the effect of migration on the migrants native areas and the left behind have originated. In fact, a lot of empirical work has been carried out using data sets from different source countries, such as Mexico and Egypt. This survey aims to provide a picture of the studies about the left behind individuals, discusses the areas of concern and works to be done in the future and concentrates on the case of rural-urban migration in China.

[3] http://www.rba.gov.au/publications/bulletin/2011/sep/4.html

1.2.1 Introduction

Antman (2013) is a survey paper about the impact of migration on the left behind family members. It reviews the literature on the effects of migration on the education and health of non-migrant children, as well as the labour supply of non-migrant spouses. It also addresses the impact of children migration on contributions toward non-migrant parents, as well as the effect on parental health. The empirical work of migrants left behind can generally be concluded into three categories - literature on the children left behind by their migrant parents, on the elderly parents left behind by their migrant children, and on the left behind siblings and spouses of the migrants.

The study of the effect of parental migration on the left behind children often evaluates these children's education outcome, health status, time allocation, etc. Parental migration affects the left behind children mainly through two channels. On the one hand, the parents migrate and become employed, which brings more income to the household and relaxes the budget constraints. On the other hand, parents' absence at home induces a lot of problems for the children, such as lack of supervision on the study, getting emotionally hurt, potential migration and crime.

The recent studies have also focused on the attention given to the left behind elderly parents. Recent empirical studies mainly examine the health status of the elderly parents of migrants. While the left-behind parents may receive remittances from their migrant children, but they may also need their children to spend more time with them. In Chinese rural-urban migration cases, elderly parents left behind in rural areas mainly depend on the financial support from their migrant children or migrants' siblings. Since the ageing population in China is steadily increasing, the subject of left behind elderly parents is a leading potential area for future studies. Furthermore, the first generation of migrants (born in the 1960s) are getting older and most of them will permanently return to their home villages, while their children would now become the new generation of rural-urban migrants (those born in the 1980s and 1990s).

The effect of migration on employment and time allocation of left behind spouses remains vague. The spouses of the migrants may need to spend more time on the house-

work. They may become more involved in farm work or the home business. Also, the education attainment and employment of siblings may be affected by the lack of labour in the family and the relaxed financial budget constraints.

The employment of migrants at the destination may be linked to the economic environment, the network, and the characteristics of the migrants. Return migrants with more physical capital, work experience and human capital are more likely to be self-employed. The employment patterns of the left behind vary between the different kinds of migrations and different destinations. Existing empirical results about the labour supply and employment pattern of the left behind household members are distinct.

Migration changes the structure of the households and may affect the resource allocated to the left behind members by affecting the bargaining power of the household members. Most empirical results indicate that parent migration, especially the father's migration favours the welfare of their daughters. Migrations affect the well-being of the left behind by affecting the households decision of consumption, physical investment, and human capital investment etc.

1.2.2 The Left-Behind Children

A majority of past literature has studied the development of children in LDCs. A lot of children are left behind in rural China when parents migrate to work in urban areas. Due to Chinese institution barriers, migrants are treated as "illegal immigrants" in urban areas of their own country. They are compelled to work long hours, doing heavy, manual work with low payment returns when compared to urban residents. Moreover, they do not have access to public welfare facilities, such as medical insurance, unemployment benefits, etc. The children of migrants cannot attend the public schools in cities without "Hukou". Most of them are left behind, while a very few migrate to study in urban private schools of lower quality.

In the context of international migration, some recent papers have been published about the left behind children. Rossi (2008) provides a survey on the impact of migra-

tion on children in developing countries. This paper discusses this impact with regards to health and education of the left behind children; children labour supply; gender differences and household allocation, etc.

Edwards and Ureta (2003) uses data from El Salvador and finds that remittances have a much larger impact on the hazard of leaving school than the household income net of remittances. Similarly, Antman (2012b) carried out a study in Mexico that evaluated the effects of parental migration on educational outcomes of the children left behind. She exploits the variation in siblings' ages at the time of the fathers migration and finds a positive effect of parental U.S. migration on the children's education attainments. It is suggested that preponing fathers migration earlier in the daughter's life can lead to an increase in her education attainment relatively by up to one year when compared to delaying the migration until after she has turned 20.

Antman (2011b) estimates that the channel through which parental migration affects left behind children includes spouse control over the intra-household allocation of resources. By comparing recent migrants and return migrants, it is found that the households that still have a head in the United States devote a lower fraction of resources to boys. It is concluded that international migration increases the decision-making power of women, while the household head is not present, and women subsequently shift resources to girls. However, once the head returns, he compensates for his absence by increasing the resources for boys.

Antman (2011a) studies how a father's U.S. migration affects the intensive and extensive margins of children's participation in schools and work in Mexico. Overall, it is found that the children reduce study hours as a result of father's U.S. migration. The study provides some evidence regarding an increase in work hours outside the home. By using children individual-level fixed effects and instrumental variables (IV) strategies, it is shown that the younger boys respond to paternal migration by decreasing their focus on school and increasing their attention on work outside the home.

Education of Left-Behind Children in China

A lot of studies have been undertaken on the education of the left behind children of the migrants in China. Wang (2014) finds a negative effect of the migrant parents on the educational attainment of left behind children and provides evidence that the negative effect is larger for boys than for girls. Furthermore, the longer the father's absence, the bigger the negative impact on boys. Meyerhoefer and Chen (2011) finds that parental migration is associated with a significant 0.7-grade level lag in the educational attainment among girls. This negative effect is not significant for boys and suggests that parents migration may lead to an increase in home production time for girls. Leng and Albert (2010) suggests that father's migration reduces the enrollment among sons, and this has a significant positive effect on the education outcomes of daughters, but has negative effects on the well-being of both boys and girls. Lu (2012) shows that migration of siblings generates benefits for children's education, while parental migration does not offer the left behind children a significant advantage in educational prospects. Moreover, young children seem to be especially susceptible to the disruptive effect of parental migration.

However, some studies have found an opposite result. Chen et al. (2009a) does not find any significant negative effect of migration on the school performance, and suggests that educational performance improves in the migrant households, where father is the one, who has migrated.

Also, some children may migrate to study in the city where their parents are employed. Chen and Feng (2013) finds that migrant students who are unable to enroll in public schools perform significantly worse than their more fortunate counterparts, and suggests that access to public schools is the key factor for determining the quality of the education that migrant children receive. Existing empirical results tend to suggest a negative or non-significant positive effect. They also show the existence of gender differences in the education outcome of left behind children. However, the long-term effect of migration on the children has still not been examined and future studies in this field should take advantage of larger sample panel data-sets.

Other Issues of Left-Behind Child

This study had also considered the time use pattern of these children. Chang et al. (2011) finds that parental migration increases the time spent on farm work and domestic work for left behind children, and this increase in work time is greater for girls than for boys. Chen (2013) finds that when a father migrates without his family, the children spend more time in household production, and this result does not appear to be driven by an increase in the mother's bargaining power.

Some studies also evaluate the health of the left behind children. Zhang (2012) finds no significant effect of father's migration on children's nutritional status, if he first migrates when the child is under the age of 6. Children are not as expected to benefit from increased household income through remittances and increased mother's barging power. Lee (2011) finds some negative impacts of parental migration on children receiving health care and on the general quality of their health. They conclude that parental absence due to labour migration affects children's mental and physical well-being.

De Brauw and Giles (2008) investigates how an increase in migration opportunities affects the decisions of middle school graduates to attend high school in China. They show a negative relationship between migrant opportunity and high school enrollment. They suggest that this relationship may be a result of an increase in subsequent local and migrant non-agricultural employment of high school-age young adults as the size of the current village migrant network increases.

However, the relationship between return decision and children left behind has rarely been investigated. Démurger and Xu (2011) contributes to this area by examining how left behind children influence the return migration in China. They find that left behind children are more likely to pull their parents back to the village, and the effect is stronger for pre-school children. When the gender difference is considered, it was concluded that the sons are found to play a more important role than daughters in reducing the migration duration of their parents.

1.2.3 The Left-Behind Spouses, Siblings, Household Bargain and Allocation

Few studies that have delved on the women's labour supply in the case of men's migration. Mu and Van de Walle (2011) evaluates how the health, labour supply and time allocation of non-migrant women in rural China is affected by migration. They find little impact on left behind women's health outcomes, but discover that women left behind are doing more farm work than the women in non-migrant household. They also suggest that there is no evidence of increased decision-making responsibilities on the left behind women over the household's farming activities. This result is contradictory to the empirical findings in other countries.

In contrast to Mu and Van de Walle (2011), Chen (2013) finds that, when the father migrates without his family, the mother spends less time in both household production and income-generating activities. Chen concludes that the burden of household production is partially shifted from mothers to children, when the father migrates.

As for the migration effect of sibling, Chiang et al. (2012) states that modest gender differences in favouring boys exists in educational migration, but there are no gender differences in the overall likelihood of labour migration. Youths with older sisters are less likely to migrate, while youths with younger brothers are more likely to migrate. It is concluded that a youth with older sisters is also negatively related to being a local or a migrant student, while for boys, labour migration may serve as a backup plan, should they fail in the high school entrance examination.

Some studies evaluate the effect of migration on the household bargain and resource allocation of the left behind. Antman (2010b) investigates how parental migration affects the allocation of resources within the household while a parent is away and after he has returned. Overall, it is identified that when a father migrates to the U.S., it decreases the fraction of expenditure on boys as compared to girls, in both education and clothing. Upon the father's return to Mexico, the fraction of expenditures on sons will rise and go beyond the initial allocation.

Regarding household allocation, it still remains unknown how the migrant households deal with the remittances in rural China. Zhu et al. (2012) state that remittances are largely used for consumption purposes by rural households, and they suggest no evidence of any direct effect of remittances on either capital input or gross output of farm production. Rozelle et al. (1999) show that the net impact of migration and remittances on maize production is negative even though the negative effect from low family labour is partly compensated by increased remittances. Chen (2006) argues that in the case of migration, the ultimate effect on intra-household allocation will depend on the capacity for monitoring and the preferences of decision makers left behind. Chen (2006) suggests that this information problem limits the allocation of remittances to easily observable goods, and non-cooperative behaviour in the household may create inefficiencies in investments and hinder growth.

The empirical results of intra-family effects of migration are disparate and typically focus on the short-term effects of migration. Future study should consider the long-term or dynamic impact of out-migration on the intra-family allocation and household decisions on saving, consumption, human capital investment, physical capital investment, farm production, etc.

1.2.4 The Left Behind Parents

The literature on the impact of migration on the left behind parents is increasing. Antman (2010a) studies, "whether elderly parents in Mexico of children in the U.S. suffer from worse health outcomes than their counterparts with no children in the U.S.". She finds that "a child's U.S. migration is associated with a greater chance that the elderly parent in Mexico will be in poor physical and mental health". It concludes that "there is a causal link pointing to worse health outcomes for elderly parents in Mexico with migrant children in U.S." Antman (2012a) considers the intra-family allocation of elderly care between siblings, when migrant children may provide financial assistance, but cannot' offer any physical care for their elderly parents. She finds that a sibling's financial contribution functions as strategic complements, while a sibling's time contributions

operate as strategic substitutes.

Guo et al. (2009) examines "the impact of out-migration of adult children on older parents' inter-generational support and psychological well-being in rural China", and discovers that "compared to parents without migrant children in 2001, the parents of migrants had significantly more monetary support, less instrumental support and a lower level of depression in 2003", but when "the effect of inter-generational support was taken into account, older parents with more migrant children tend to have significantly more depression and lower life satisfaction."

Giles et al. (2010) finds that in Chinese rural-urban migration cases, elderly parents with migrant children will receive similar levels of financial transfers, as those without migrant children. They argue that the predicted variances associated with these transfers imply that there is a high risk that elderly parents who have migrant children could fall into poverty.

Chang et al. (2011) studied the impact of migration on time use patterns of the left behind elderly parents in rural China and finds that migration will lead to an increase in the time spent on farm work and domestic work for the left behind elderly. Significantly, the effect of increase in work time is greater for elderly women than elderly man.

However, Zhongdong and Guowei (2009) argues that the impact of migration on elderly care is not often negative, and a loss of care can be compensated by remittances from migrant children to elderly parents. They show that the left behind elderly parents with inter-provincial migrant children are found to be more economically satisfied than those in a household with no migrant children. Besides they provide evidence that the negative and positive effects of temporary migration are more or less balanced as the difference of the overall well-being of elderly parents with migrant children and those without is minimal.

Contrary to the studies evaluating the impact of migration on left behind parents, Giles and Mu (2007) shows how the migration decisions of adult children in rural China are affected by the health conditions of their elderly parents and the elderly care requirements. They discover that "younger adults are less likely to migrate when a parent is ill";

however, when the migrants have siblings left behind to provide elderly care, they show that "poor health of parents has less impact on the probability of children's migration decision".

As the population in China, as well as in many LDCs, is ageing fast, the research on the left behind elderly parents seems growing. Among the existing empirical works, most examine the relationship between childrens migration and parents economical satisfaction and health conditions. To the best of the knowledge, there is no empirical work on the left behind parents and children's return decision except Démurger and Xu (2013), which focuses on the left behind children and the return decisions of migrants. Research utilising effective panel data may have significant policy implications.

1.2.5 Migration, Network, Rural Labour Market and Employment of the Left Behind

Past researches on the migrating decisions are primarily focused on the network effect. McKenzie and Rapoport (2007) suggests that as community migration networks grow, wealth becomes less of a constraint on individual migration from Mexico to the U.S., and the poor are more likely to migrate. It is identified that large networks spread the benefits of migration to members at the lower end of the consumption and wealth distribution of the community.

Zhao (2003) studies the role of migrant networks in labour migration in the case of rural to urban migration. It is shown that the experienced migrants have a significant positive effect on subsequent migration, but return migrants do not. She suggests that migrant networks are important and their effects materialise through practical assistance in the migration process. Chen et al. (2010) examines the peer effect in Chinese rural to urban migration and finds a 10% increase in the migration rate of co-villages raises one's migration probability by 7.27%, and it is estimated that this effect is comparable to an increase in education by 7-8 years. They also provide evidence that states that most of this effect is driven by co-villages helping each other in respect of moving costs

and job searches.

Furthermore, Zhang and Zhao (2015) finds that there is a relationship between social-family networks and self-employment of temporary rural to urban migration in China. It is suggested that rural-urban migrants with larger social-family networks are more likely to be self-employed in the city. While, Parodi et al. (2012) state that the self-employment of migrants is positively selected in terms of their unobservable characteristics, and wage differential is found to be an important determinant of self-employment.

Some studies have examined the rural employment and patterns of labour supply of the left behind. Démurger and Li (2013) shows that individual migration experience favours subsequent local off-farm work. At the family level, migration pushes the left-behinds to farming rather than to off-farm activities. This is in line with Giulietti et al. (2013), which finds that return migration promotes self-employment, while the left behind individuals in rural areas are less likely to be self-employed. It is suggested that migration may have a negative impact on the self-employment of the left behind because of the absence of the migrant.

The impact of migration on the self-employment of return migrants is positive. McCormick and Wahba (2001) finds that both overseas saving, and the duration of a stay overseas, increases the probability of a return migrant becoming an entrepreneur, among literate returnees to Egypt. While amongst illiterate returnees, overseas savings alone increases the probability of becoming an entrepreneur. Ma (2001) finds that the improvement of the migrants' skills and entrepreneurial abilities, rather than their savings and remittances, strongly facilitates an occupational change in rural China. Ma (2002) studies the mobilisation of social capital - the channel through which migration experience affects entrepreneurship. Ma argues that human-capital accretion during the migration reinforces the mobilisation of local social capital after return and which, in turn, enhances income return to rural entrepreneurship. Démurger and Xu (2011) analyses the self-employment decision of migrants and finds that both return savings and job change frequency increases the likelihood of the return migrants becoming self-employed.

1.2.6 Discussion

The literature regarding the left-behind children and elderly parents of the migrants is ever-growing and developing. Although a lot of empirical work has been conducted on the health and welfare of the young and old, as well as on the education of the young in rural China, the results vary and future work is desired to produce exact and explicit cause-effect results.

Considering the scarcity of studies, the determinants of migration or return decision, need to be investigated. The investigation of left behind spouses is a topic of interest as it involves the household bargaining and household allocation, which may affect the consumption behaviour, investment decisions, children outcome and production of the household.

The externality of return migration is critical to the rural household, as well as for the development of rural areas. The return of migrants with savings, skills and knowledge not only affects the employment pattern of the migrants but also influences the left behind family members and other families in the villages. Existing studies regarding the employment change of return migrants almost all conform to each other, however, the impact on the employment, labour supply and welfare of the left behind non-migrants is far from substantiated. Besides, there are limited studies that relate the return migrants with the rural development in China, although there are numerous studies that focus on the return migrants in Mexico, Egypt and Spain.

1.3 Survey on Migration, Norms and Behaviour

The impact of return migration on the originating areas or countries is a growing area of interest in the migration studies (see Wahba (2014) for a review of the economics literature on the impact of return migration). As mentioned in the previous survey, literature on the migration or return migration typically does not emphasize on the migration sending area. Despite the existing and growing empirical or theoretical work on

the investment, saving and labour market behaviour of return migrants, literature on the behaviour of returnees is exiguous.

The theory of economics migration predicts that migrants would accumulate assets and human capital in migration destinations, and then finance higher consumption, or change their investment and saving behaviour upon return (Dustmann and Görlach (2016)). The mobility of the population in different places of habitat not only impacts the financial and human capital of the migrants, but it also affects the other outcomes. Largely, most studies on the impact of migration focus on the economic returns of migration, while ignoring the social returns, especially the externalities of migration and return migration. Migration influences the norms, knowledge and behaviour of the migrants, thus the new knowledge and ideas that the migrants have been exposed to would be transmitted to the origin of migration, as well as residents living there.

1.3.1 Introduction

Migration influences the migrants and sending areas in many areas (see Wahba (2014), Dustmann and Görlach (2016) for reviews). There is some emerging literature investigating the impact of return migration on the transmission of social norms. For instance, Bertoli and Marchetta (2015) shows that Egyptian return migrants adjust their fertility choices to the norms in migration destination countries; return migrants, who have past migration experience in another Arab country, have a significantly larger number of children as compared to the stayers. Similarly, in the analysis of this thesis, migrants adapt to the urban life, and the migrants returning to their home village may change their behaviour according to the norms that prevail in the cities, where they worked. They also bring back social norms and knowledge they experienced and gathered in the migration destination.

Some papers study the impact of migration and return migration in various aspects. For instance, Batista and Vicente (2011) studies the impact of migration on political ideas and Wahba and Zenou (2012) studies the impact of migration on entrepreneurship.

Specifically, migration could change the behaviour and beliefs of individuals who have migration experience. The literature on the impact of migration on social norms, behaviour change in economics is limited but emerging. For instance, Tuccio and Wahba (2015) investigates how international migration transfers gender norms in Jordan. Similarly, Bertoli and Marchetta (2015) studies the impact of return migration on the fertility choice of Egyptian women.

However, there is no relevant study in this area in the context of rural to urban migration in China. Certain studies, such as Chen et al. (2009b), Yang et al. (2007), in the field of public health and medicine have examined the impact of migration on health behaviours. The analysis in the current thesis would be the first study about China in this field.

1.3.2 Migration, Norms and Behaviour

By contributing to this emerging field, the current thesis includes two papers that study the impact of migration and return migration. The third chapter analyses the consumption behaviour change of return migrants, while the fourth chapter evaluates the impact of migration on diet knowledge. Migration experience would impact the migrants, and the two chapters in this thesis will try to find out whether migration has a spillover effect. In particular, they would focus on the norms of return migrants and whether migration spread knowledge to other residents in the village.

It is customary to study the diffusion and spillover effect in the economics literature, focusing especially on technology, innovation and growth (see Jaffe (1986), Barro and Sala-i Martin (1997) for an earlier study, and a review for international technology diffusion Keller (2004)). Recent studies include Bloom et al. (2013) who look for a technology spillover effect using U.S. firm panel data and Aghion and Jaravel (2015), who study the R& D and knowledge spillover in the economic growth process. Apart from these, some papers study the spillover effect in the development aspect. Spolaore and Wacziarg (2009) studies the diffusion of development from the world technological

frontier, and addresses the importance of the genetic distance as a barrier to diffusion. Banerjee et al. (2013) examines how participation in a micro-finance loan programme diffuses via social networks in Indian villages. They examine the mechanism of diffusion taking advantage of the process of information transmission in social networks.

There are some relevant studies regarding the process of the spillover effect of migration. Kerr (2008) studies how U.S. high-skilled migrants help in the process of technology diffusion to their home countries, through their ethnic networks. Similarly, Hornung (2014) studies how Huguenot migrants in Prussia help to diffuse technology across the borders. These papers study the spillover effect of migration on technology and productivity in the context of international migration. Compared with these papers, this thesis centres on the internal rural to urban migration in the development context.

1.3.3 Discussion

The study of the social-political impact of migration and return migration is emerging but still limited. As pointed out by Wahba (2014), the research on the behaviour and knowledge of migrants and return migrants are constrained by good quality data. Future research agenda would have plans on the survey design and data collection for the study on the impact of migration and return migration.

The other point this thesis aims to address is that policy makers should have a proper policy to maximise the benefits of migration and return migration. This is especially true in the case of Chinese policy makers, where they still have the "Hukou" system. Moreover, they should not only consider the economic return of migration for individual migrants and the economy of the destination, but also the social return of migration, which is important for developing areas.

1.4 Data: China Health and Nutrition Survey

The China Health and Nutrition Survey (CHNS) is an international collaborative project conducted by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention. The survey was conducted by an international team of researchers and experts in the field of nutrition, public health, economics, sociology and demography. The survey has 9 waves sample from 1989 to 2011 (still ongoing for the 2015 wave), including about $4,400$ households with $26,000$ individuals across 9 provinces, which vary substantially in geography, economic development, public resources and health indicators. Detailed community data was collected in surveys of food markets, health facilities, family planning officials, and other social services and community leaders. The three chapters of analysis in this thesis take advantage of the abundant multilevel social and economic statistics.

1.4.1 Survey Design

The study population of the survey is randomly drawn from the provinces of Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong. Most of these provinces are traditionally inland agricultural provinces. As described in the official documentation of CHNS, "A multistage, random cluster process was used to draw the samples surveyed in each of the provinces. Counties in the nine provinces were stratified by income (low, middle, and high), and a weighted sampling scheme was used to randomly select four counties in each province. In addition, the provincial capital and a lower income city were selected when feasible. Villages and townships within the counties and urban and suburban neighbourhoods within the cities were selected randomly." The follow-up levels are high in the data, but families that migrate from one community to a new one are not tracked. The tracing of the migration destination is not available in the data, which limits the migration study of this thesis. The research for this thesis focuses on the rural areas and individuals in the rural villages. As the

counties and the villages within the countries are selected randomly, the study samples in this thesis are considerably random, and picked from nine provinces.

Since 1997, new households in original communities were added to replace the households that are no longer participating in the study. Also, since 1997, new communities in the original provinces have been added to replace sites that are no longer contributing. A new province was also added in 1997 when one province was unable to participate. The analysis in this thesis uses the 1997 to 2011 waves (in which migration information is available) for the first and second paper and 2004 to 2011 waves for the third (in which diet knowledge data is available). Thus, the significant attrition that occurred before the wave of 1997 does not impact the analysis in this thesis.

1.4.2 Weighting

It is important to use the sample adjusted by sampling weight for analysing with survey data. However, the sampling weight for this data is not available. In the official documentation of CHNS, it is stated that "when the survey was planned and implemented, the State Statistical Office of China would not share their sample frame with the CHNS team. Furthermore, we could not have released the data sets for public distribution if we had worked with them. The design used extant census data as best as we could for a multi-level random sample. However, it was determined by our sampling statistical colleagues that we could not create even cross-sectional sampling weights, let alone longitudinal ones."

This thesis elects not adjusting the sample by weighing for two primary reasons. First, the data is collected by a multistage random cluster process, which would make the sample random and thus representative. Also, it is not possible to obtain the sample frame from the national statistical office of China. Second, this thesis uses the various levels of control, as described and recommended by the CHNS team, "to control for multilevel multistage sampling and an array of multilevel modeling issues, we recommend that researcher utilises various levels of control. These include community

level measures of the newly created urbanisation index which is now (in the year 2014) available with the CHNS Longitudinal Data."

Furthermore, this thesis utilises the most updated data of CHNS and controls many relevant characteristics at different levels - individual, household and village level available measures. Certain analysis in this thesis also use the confidential community survey. Moreover, this thesis applies robust and cluster-robust standard errors, wherever available, in order to make the results more convincible and reliable.

1.4.3 Data Collection

This part describes the details regarding the information collected in the data. The information that is relevant to this thesis is addressed below.

The survey contains detailed information on demographic, economic, time use, labour force participation, and asset ownership aspects. The first analysis of time use is based on this information. The income records are detailed; "income can be approximated from the survey in three ways: through responses to direct questions about income, the summation of net receipts from all reported activities. Full income from the market and non-market activities can be imputed. This detailed estimation of income represents a significant advance in the measurement of income in China. The inclusion of non-monetary government subsidies, such as state-subsidized housing, is an especially important advance" (CHNS official release).

The health survey contains information of individual data on dietary intake, body composition, and health-related behaviours and knowledge (e.g. smoking, beverage consumption, etc.). The data included clinical measures of health; and measurement of weight, height, arm circumference, and head circumference. This information is used in the second and third papers. In the health section of the survey, details about insurance coverage, availability of medical facilities and illness information are covered. This information is used as controlled variables in the relevant research analysis.

The community questionnaire (filled out for each of the primary sampling units)

collected information from a knowledgeable respondent on community infrastructure (water, transport, electricity, communications, and so on), services (family planning, health facilities, and retail outlets), population, prevailing wages, and other associated variables.

The CHNS claims that all field workers are educated to a certain level and training has been provided to the field workers. Thus, the data collected can be considered to be reliable and trusted for this thesis study.

1.5 Methodology

The self-selection of migration is the primary challenge for identification in migration studies. There can be various econometric methods in migration studies that depend on the data, as well as on the research questions. To determine a causal effect of migration and return migration is the chief theme of this thesis. Thus, it's important to identify econometric models to deal with the selection and endogeneity issues in the analysis.

The fixed effect models can be applied to deal with the unobserved time-invariant that may lead to the endogeneity issues. The panel nature of the data applied in the thesis provides the fixed effect method to control for the time-invariant sources of endogeneity. All the three analysis of the thesis applied the fixed effect models. Depending on the research questions, the three papers apply individual fixed effects in order to eliminate the individual time-invariant unobserved factors that lead to the endogeneity issue, or the selection of migration and exposure to mass media. The common aim of the fixed effects models is to allow for the arbitrary correlation of the observed or unobserved time-invariant variables and the interested endogenous regressor.

The other methodology that is usually applied in migration studies is the instrumental variable (IV), which is applied to deal with the selection issue, the time variant endogeneity source of migration or mass media in the models. In the studies on migration's impact on the left behind, some use the IV approach and proposed instrument variables.

On the other hand, certain studies use the historical migration rate in sending regions and the migration stock in the destination region, as the migrant network (Mansuri (2006), Hildebrandt et al. (2005), McKenzie and Rapoport (2007), Binzel and Assaad (2011), Lokshin and Glinskaya (2009), Grigorian and Melkonyan (2011)). Instrument variables may also result from an understanding of the economic phenomena. Chen et al. (2010) uses the gender of the neighbours' first-born and the sex ratio of labourers in the neighbourhood households as an instrument for migration. Amuedo-Dorantes and Pozo (2006) instrument remittances with information from the West Union, an American financial services and communications company. Grigorian and Melkonyan (2011) uses a measure of corruption as an instrument variable for migration. Some papers on migration in China also choose instrument migration. For example, Chang et al. (2011) uses the village level migration rates as an instrument for migration.

Some have tried to use the economic situation in the migration destination as an instrument for migration. McKenzie and Rapoport (2007) uses the unemployment rate in the US as an instrument for migration from Mexico to examine the dynamic impact of migration on inequality. Antman (2011a) uses the employment rates in popular US migration cities to instrument for the migration from Mexico in order to study the impact of migration on left behind children's schooling.

Similar to McKenzie and Rapoport (2007) and Antman (2011a), the first paper of this thesis uses the economic conditions in the migration destination as an IV. The first paper of this thesis uses the wage gap between potential destinations and local cities around the villages. Employment and unemployment rates are not used, as employment statistics recorded in Chinese cities only record those city locals with urban "Hukou", who report their employment statuses, and in general there is little variation in employment statistics across cities and over the years.

The second analysis uses another novel IV, which is the individual body mass index, different from the average of peers. The property of employment and industry cluster of migrant workers suggest that taller and healthier individuals from the rural areas are more likely to be employed in the coastal urban areas. This shows a positive correla-

tion between the BMI and the incidence of migration choice. Furthermore, the current consumption behaviour is not correlated with the body status of the past (the last wave).

The last analysis partakes no issue of migration selection as the interested variable of the migration density at the village level. However, the individual mass media exposure is endogenous. The village level mass media exposure is applied to instrument the individual mass media. The correlation is clear; village mass media impacts the individual knowledge through two channels, namely the individual mass media, or the village knowledge. The village knowledge is captured by the peer knowledge effect in the model, thus the IV can only impact the individual knowledge through the endogenous variable, which is the individual mass media.

This thesis tries three different IVs based on the different research questions, contributing to the area on causal effect identification. The amalgamation of the fixed effect and IV method in the econometric modelling helps to deal with the self-selection and any endogeneity issues in the exercises. Moreover, the inclusion of time fixed effects with FEIV would provide a proper estimator.

1.6 Links of the Papers and Contribution

The three papers of this thesis contribute to the economics of migration that emphasizes the outcomes in the migration sending areas; the left behind family members, the return migrants and the non-migrant rural residents. The different studying objects in the three analysis of the thesis are the three important elements in the developing areas of China, and they are also the less studied objects in the academic and policy analysis.

The overall theme of this thesis is understanding the impact of migration on the rural residents and rural areas. The analysis begins from the time allocation of the left behind family members. This is an under-studied topic in the academic research. The left behind family members, especially the children, have been the subject of public discussion and draw attention from the policy makers. For instance, the extreme poverty living

conditions, result in crime and some suicide cases amongst the left behind children [4]. This thesis focuses on the time use of the left behind children and spouse of the male household head. The extended family, or parents of the household head, are normally left behind as well. There is some interest in studying the outcomes of the left behind parents of the migrants, such as their physical health and mental health. However, the majority of the left behind family members and the core family members are studied in this thesis.

The family members of the migrants are left behind in the rural areas, mainly because of the "household registration" policy in China. This policy normally identifies the individual's status (rural or urban registration) by their place of birth or the status of their parents or spouse. This policy has been very strict for many years and has had a huge impact on rural residents. The household registration system "creates many difficulties for migrants living in urban areas in terms of good and secure jobs, housing and access to public services and these difficulties deter or prevent migrant workers from bringing their families with them to urban areas" (Knight et al. (2011)). The migrants are normally not left with a choice, but to leave the family behind in rural areas, and temporarily migrate before returning to home villages.

This introduces the topic of the second analysis, which focuses on the return migrants. Migrants in China can rarely bring their family with them or settle in the destination. The rural to urban migration in China is temporary and circular. The return migrants bring not only the financial and human capital back to their village upon returning but also the behaviour and social norms they experienced and adapted themselves to in the migration destination.

Behaviours, such as smoking and drinking, are given a special attention. They are not only consumption behaviours, but also addictive habits that can be harmful to the migrants themselves and also for others, such as family members, friends, relatives, etc. More importantly, smoking and drinking are popular consumption habits in China

[4] http://www.economist.com/news/leaders/21674715-there-are-70m-reasons-ease-chinas-curbs-internal-migration-pity-children

and cause many problems in public health domain. Motivated by these aspects, the second paper of this thesis focuses on the smoking and drinking behaviour of the return migrants.

The study on the impact of migration on social norms and behaviour is emerging in the context of international migration. However, the analysis of consumption behaviour of return migrants in this thesis is the first paper in the field. It is found that the return migrants are more likely to smoke and drink, and in higher frequency, as compared to the rural residents who do not have any migration experience. This finding is consistent with the literature in the empirical study of international migration.

This result motivates to further analyse certain questions like, would migration density of the villages have an impact on the rural residents, and if there are any externalities of migration in rural China. Hence, the third analysis evaluates the spill-over effect of migration. Given that migration and mass media are the two most important channels for rural residents in China to get information, the role of mass media is also considered. It is expected that the migrants who are exposed to diet knowledge would bring their knowledge about healthy diet back to the village. Therefore, a higher migration density would increase the diet knowledge of the rural residents.

To conclude, the three papers in this thesis are logically related and contribute to the existing migration studies. Although the three chapters focus on different objects and outcomes, they gradually develop on the basis of the main theme - the impact of migration on the rural residents in Chinese rural areas.

2

Migration and Left Behind

2.1 Introduction

China has seen dramatic changes in its labour market over the past few decades, with hundreds of millions of working-age peasants moving from rural to urban areas (Démurger and Li (2013). This is the largest labour flow in world history (Zhao (1999); Knight and Gunatilaka (2010); Knight et al. (2011)). The thing is rural migrants working in China have been regarded as cheap labour and for a long period of time, they have been the power behind China's industrialisation and economic growth. Migrant-sending areas have also benefited from the relationship between the migrants and their non-migrant family members left behind. This benefit comes through remittances sent by the migrants, capital and human capital of return migrants and increased opportunities for urban employment based on the migrant network.

Migrants contribute greatly to rural and urban development in China when they try to

maximise the welfare for themselves and their families. Academic research has focused on the migrants living in urban areas, determinants of migration and return migration (Zhang et al. (2010); Frijters et al. (2011); Knight and Song (2003); Giulietti et al. (2013)). However, as pointed out by Démurger and Li (2013), "there is still little empirical evidence on how migration and remittances affect individuals and households who stay behind."

Many women, children and elderly parents in migrant families are left behind in rural areas, due to the multiple reasons. First, women, especially the elderly and married women, are at a disadvantage when it comes to employment in urban areas, because of the social-cultural traditions and employers' preferences for young and single women, when hiring females (Fan (2003)). Second, institutional constraints like the household registration system "create many difficulties for migrants living in urban areas in terms of good and secure jobs, housing and access to public services and these difficulties deter or prevent migrant workers from bringing their families with them to urban areas" (Knight et al. (2011)). Finally, under the land ownership policy in China, each rural household has the right to cultivate a piece of land (farmland, woodland, lakes, etc.) and labour is needed to work on this land if the households want to sell or consume the agriculture products.

Migration forces the left behind members to reallocate time across different activities - off-farm, farm work and domestic work. How migration affects the time use and labour supply of the left behind spouse and children remains an open empirical question.

Theoretically, "the global net effect of migration on the diversification of income-generating activities in sending communities is uncertain"(Démurger and Li (2013)). This paper studies the impact of male household head migration on the time use in housework, agriculture and wage labour of left behind children and women, while Démurger and Li (2013) evaluates the similar outcomes of all left behind family members as a whole using a cross-sectional data.

Migration affects the time use and labour response of the left behind through two channels, the "income effect" and the "substitution effect". In the context of migra-

tion, "income effect" is the effect of non-labour income (remittances) from migration that would raise the reservation wage of the left behind family members. The income effect of migration occurs if the left behind family members receive remittances from the migrants. "Substitution effect" comes from the fact of time-allocation for family activities like housework, household farm, etc. When one family member of the household migrates, the left behind ones need to adjust their time allocation accordingly to compensate for the lost labour.

The income effect stemming from remittances "would raise the reservation wage of non-migrants and thus potentially decrease labour force participation"(Antman (2013)). In the case of Chinese rural to urban migration, the "remittances from migrants affords the left behind household members to spend less time on off-farm or farm work"(Chang et al. (2011)). On the other hand, as migrants leave, the total labour supply of the household decreases; as a consequence, migration may increase the time and labour supply of the left behind members to compensate for the lost labour.

The income effect and substitution effect of migration work in opposite directions, the net effect of migration on the time use and labour of the left behind is not evident. In accordance with economics theory of migration and previous empirical findings, it is assumed that, firstly through the income effect, time use and labour supply of the left behind decreases; secondly through the substitution effect, left behind women and children increase their time use and labour supply. This paper focuses on the net effect of migration to ascertain whether the income effect or substitution effect has a greater impact. However, how the time use aspects of the left behind are impacted by the migration of the male household head does not reflect if a household with a male migrant household head is better off or not.

This paper investigates the time use and labour supply of the left behind members in response to migration in China, using the China Health and Nutrition Survey (CHNS), which contains time use information, migration and other social-economic characteristics of individuals, households and villages. This paper uses an instrumental variable approach to control for the endogeneity issue in the model. Lack of remittance infor-

mation in the data makes it difficult to measure the non-labour income and the timing of "income effect". This paper assumes that all the left behind family members receive remittances. It identifies that, overall, the time use and labour supply of the left behind spouse and children are not statistically different from the members in non-migrant households. The findings suggest that the income effect is offset by the substitution effect, while the net effect of migration is not considerably significant. This result also survives various robustness checks.

Limited studies focus on the rural to urban migration in China, especially on the left behind. Research on the left behind family members are important as these are closely related to poverty reduction, inequality and rural development due to the huge number of children and spouses left behind in the vast rural areas in modern China. As estimated from the Chinese census, there were 61 million children left behind in 2010.[1] These numbers are large and are nearly equivalent to the entire population of the UK. It is estimated that, in 2006, there were around 50 million left behind women (Zhang (2006)) living in villages and separated from their husbands. At the micro level, the time use of the left behind children and spouses, are related to many other aspects of welfare for individuals and households. For instance, the education of the children may be affected. Some children may have to quit school at an earlier age to compensate for the lost labour in the family due to the substitution effect of migration. The left behind women, who are living without the male household head, may take up the main responsibilities for the household and may suffer from possible physical or psychological health issues. The migration of the male household head also relates with the household bargaining. At the macro level, the Chinese authorities would care more about the left behind families and amend their policies, as otherwise, there may be many potential social and economic problems. First, as an effect of the one-child policy, after decades, the population size of the younger generation is rapidly decreasing (Figure 1.5). The well-being of younger generation is critical for the country's future. The health, education and emotional situation of the left behind children would have a huge impact

[1] Chinese the official news agency, The people's daily http://acwf.people.com.cn/n/2013/0510/c99013-21437965.html

on the level of human capital of the country. The left behind children, who receive less supervision from their parents are more likely to adopt behaviours like smoking or drinking, and may even incline towards criminal tendencies. The left behind family, living separated, would also experience many other problems, and may lead to certain questions like, would the left behind women be able to maintain the agricultural productivity; would a lot of farm land be abandoned due to the lack of labour; and, would the left behind women choose to decrease the fertility rate. All these issues lead the policymakers, the public and academia to think more about the left behind children and women.

Other sections of the chapter are structured in the following pattern. Related studies are reviewed in section 2, while section 3 introduces the data and presents the descriptive statistics; section 4 shows the model and the instrumental variable, followed by section 5 that presents and interprets the main results; section 6 conducts robustness checks and section 7 offers a conclusion.

2.2 Literature Review

Existing articles about the impact of migration on the time use of left behinds are limited. Moreover, most of these studies focus on the labour supply of the spouse or the employment patterns of the left behind family members, and only a few shed light on the time use and labour response of the children left behind.

In the context of international migration, some papers have conducted empirical research in several migrations sending countries. The primary finding is that left-behinds in the international migration households reduce their labour supply (Kim (2007) for Jamaica; Rodriguez and Tiongson (2001) for the Philippines; Grigorian and Melkonyan (2011) for Armenia, etc.). The issue is that they do not specify the role of the left behind in households.

Also, empirical findings in some developing countries are ambiguous. Cabegin

(2006) examines the effect of one partner's international migration on the other non-migrant partner's labour supply in the Philippines, and the results present a stronger substitution effect of migration for married women and larger income effect for married men. Amuedo-Dorantes and Pozo (2006) concludes that it is difficult to disentangle the income effect and lost labour effect, as remittances may increase or reduce the working hours, depending on gender, location and types of employment. Görlich et al. (2007) finds that migration in Moldova increases the likelihood of engaging in home production for the left behind households, while the labour market inactivity of the left behind households is not due to leisure, but because of the participation in higher education and home production. Hanson (2007) states that in Mexico, individuals are less likely to participate in the labour market if the households send international migrants or receive remittances.

Some studies focus more specifically on the spouse left behind. Binzel and Assaad (2011) finds that in Egypt, the migration of males decreases the wage work of women in urban areas, while rural women do more unpaid family work and subsistence work. Lokshin and Glinskaya (2009) finds a negative effect of male migration on the left behind women's wage labour. Hadi (2001) does not find a positive effect of the husband's international migration on the left behind wife's labour market participation in Bangladesh and recommends further study on the effect of a spouse's migration on the different occupational outcomes for women.

For the left behind children, most studies focus on the health and education outcomes, although a few focus on the child labour participation and time use. Antman (2011a) studies the effect of paternal American migration on the schooling and work of children left behind in Mexico and concludes that children reduce their study hours and increase work hours when the father migrates to the USA, and the effect is stronger for 12 to 15 years old boys. Using data from El Salvador, Acosta (2011) finds a strong reduction of child wage labour in remittance-receiving households. These two papers investigate the labour supply of children left behind.

Chinese studies on the left behind in the context of rural to urban migration is lim-

ited. De Brauw and Giles (2012) finds that migration increases the labour supplied to productive activities for poorer households in China. Démurger and Li (2013) considers individual occupation change across farm work, wage work, self-employment and housework for the left behind.

Articles closely related to this paper are as follows: Chang et al. (2011), using the same data as this paper, studies the time use patterns of the left behind children and elderly, and finds that rural to urban migration increases the time spent on the farm and domestic work for the left behind elderly, as well as children. Mu and Van de Walle (2011) finds that left behind women do more farm work when considering the women's participation and time allocation in different types of work. They also find that the presence of children in the household would change the time use and labour supply of the left behind women or men. Chen (2006) and Chen (2013) both find that left behind children spent more time in household production, while the left behind spouse spent less time in both household production and income-generating activities.

Although using the same data as the above studies, this paper differs from the literature in the following aspects. Firstly, the newly available village data from the updated database would add many extra village level controls that account for potentially omitted variable problems. Secondly, this paper innovates in econometric methodology by using the panel property of the data, as well as dealing with the endogeneity issues of migration. Chen (2006) and Chen (2013) does not consider the endogeneity property of the migration and applies the simple regression models. Chang et al. (2011) uses the pooled cross-sectional data and the migration rates in the village as instrumental variables, which are supposed to impact the labour supply of the left behind. This is because the migration in the village would change the total available labour in the village. Mu and Van de Walle (2011) uses the first difference, combing lagged controls as the proposed instrumental variable, which is not very convincing, as the controls are auto-correlated. Thirdly, this paper offers a more detailed study outcome in an extensive and intensive measure for both the left behind children and spouse, which are not just the core and most influenced left behind family members, but also the majority of left behind members.

Given the ambiguous existing findings, this paper uses a household survey to study the income and substitution effects of rural to urban migration in China. The probability of participation in, and the hours per day spent on, such activities as agriculture, household work, and wage labour are examined for both left behind spouse and children.

The aim of this paper is to contribute to the limited empirical studies on how migration impacts the left behind, particularly the time use and labour supply of the left behind children and spouse, by settling down the total effect of migration on the left behind via the income and substitution effects.

2.3 Data

This paper uses the China Health and Nutrition Survey (CHNS), which is conducted by the Carolina Population Center at the University of North Carolina and Chinese Center for Disease Control and Prevention. The data includes 9 waves, from 1989 to 2011 and incorporates a sample of about $4,400$ households with $26,000$ individuals, across 9 provinces in China.

The panel survey is comprised of a household survey, individual survey, nutrition & physical examination and a village survey. The household survey covers family members, their relationships, assets and properties owned by the household. There is information about individual participation and time allocation on different activities such as household farming, household chores and wage labour.

Migration is defined as "not living at home currently" and "sought employment elsewhere". The 1997 and the following waves are used. This is because, since 1997, questions about migration are included in the survey questionnaires.

This survey does not record the migration history or the destinations of migration. The migration in this paper may be underestimated, as rural to urban migration in China is circular and temporary. According to Sun and Fan (2011), nearly 86.8% of inter-provincial migrants are temporary, and 94.1% of rural to urban migrants are also tem-

porary. Considering the wave gaps in the survey are 3 and 4 years, migrants may have experienced repeated migration and could have returned during the two panels. However, this paper evaluates the short-term impact of migration on the left behind and compares the outcomes in migrant households and non-migrant households by the current migration status of the male household head. Some may argue that a migration period of a short or long duration may have different effects on the left behind. Though this can be true, however, this is another research topic that will use duration data. Duration data for migration studies in China is rare, thus the analysis of migration duration and the outcomes of left behind family members can be an innovative research in this area in the future.

The other limitation of the data is the lack-of-information regarding remittances. Thus, the amount of non-labour income the left behind families received and the timing of the "income effect" of migration is not available for this research. This paper assumes that Chinese rural left behind family always receive remittances, thus the "income effect" of migration exists and influences the time use of the left behind family members. This assumption is in line with Chang et al. (2011), Mu and Van de Walle (2011) and other literature in the field and can be verified, as it is stated in Mu and Van de Walle (2011) that "rural migrants in China typically send back remittances." Though it is a strong assumption, but it is relatively reasonable in the Chinese context as the bias is potentially small. Firstly, rural China is a traditional society with strong family ties, and it's not surprising to see the high remittances rates. As the above papers mentioned, Chinese migrants always send remittances. Démurger and Li (2013) shows that around 80% migrant households receive remittances. Second, some actual remittances can not be captured in survey data (for instance, many migrants have difficulties to use the banking facilities; many migrants are paid at the end of their temporary employment. some family receive remittances after the survey interview, or the common case that migration generated income goes directly to schools (for child education)). Finally, assuming remittances for the few non-remittances receiving migrant households may have a slight underestimation of the income effect of migration. However, this would not lead to a serious bias of the findings in this paper, as the remittance-receiving migrant households

are the dominating majority.

The left behinds are defined as the non-migrant children and spouses of the migrants. In this analysis, rural households that comprise a male household head and spouse with at least one child (where spouse and children are living in a rural area, while the male household head can be a current migrant or not) are considered for the sample. Rural households with a female household head who migrates are excluded in the comparison group. Also, the rural households with no children are also excluded from the study. This is because these households are the minority and, hence, not representative. The spouse is referred to as the wife of the male household head and mother of the children, as in most cases, the males migrate and leave the spouse and children behind in villages. This is supported by the data, which states that in 60% of all migrant families, the father migrates solely; in 25% the father and mother migrate together and, in 15%, the mother migrates alone.

Individual participation and time use information is distributed into 3 groups, agriculture, wage labour and household work. Agriculture activities include farming, fishing, gardening and raising livestock. Time spent on agriculture is the sum of time spent on all the agricultural activities combined. The indicator of participation in agriculture is 1, if an individual is involved in at least one of these agricultural activities, otherwise, it is 0. Information on wage labour come from the employment records of the individuals. Participation in wage labour is indicated by 1 if the individual is employed and gets paid for the survey year, otherwise, it is 0. Hours in wage labour are the sum of hours in the primary and secondary jobs. The paid wage work can be the employed work in agriculture or other jobs. Agriculture activities are work for family agriculture for self-consuming or market sell, and wages are not paid for agriculture activities. Household work includes buying and cooking food, washing and cleaning, as well as caring for young children and elderly parents. Participation in household work is 1 if the individual is involved in at least one household activity, otherwise, it is 0. An individual's hours per day spent on housework is a sum of the time spent on each household activity.

Migration information on rural residents is presented in Table 2.1. As the migration

rates indicate, rural to urban migration in China is extensive and growing over the years. Since 2004, the rural to urban migration of adult men has accounted for around 20% or more of the total rural male adult labour force, while for adult females the percentage is over 10%. The migration rates for the three groups have been increasing over the years. This is aside of the decrease in 2009, which shows the impact of the global financial crisis in 2008 on the employment of Chinese peasants in urban areas. Zhao (1999) pointed out that, in 1995, 8.7% of all labourers migrated, based on a rural household survey in Sichuan province. This paper observes a similar, but a lower rate than Zhao (1999) due to the fact that Sichuan is one of the largest migration-sending provinces (Figure 1.10). Furthermore, child labour is one of the primary focuses of this paper, as it is not only related to child development, health and education, but it is also closely associated with the future of the society. Child migration is not common, but it has grown significantly from 3% in 1997 to 12% in 2006. The decline in child migration in 2009 is massive. From 2006 to 2008, the nine years of compulsory education in all primary and middle schools was made free and some financial subsidies from the government were provided to the rural schools. [2] This may have motivated some rural children, who would have migrated, to stay in school and complete their education. This accounts for the sharp decline in children migration. The gender difference in migration is also noticeable and the migration rate of rural men is almost double the rate of women over the years. From the migration statistics presented in Table 2.1, it is reasonable to infer that women are more likely to stay behind with children when the male household head migrates.

[2] State Council of People's Republic of China, www.gov.cn, State Council Policy No.43, 2005

Table 2.1: Migration rate by age and gender over the years

	1997	2000	2004	2006	2009	2011
Adult man	7.51	11.29	19.38	24.75	23.72	25.97
Adult women	5.00	6.90	10.77	14.23	14.42	13.84
Children	3.26	5.93	11.36	12.21	7.67	6.22

CHNS 1997-2011. Children 6 to 18 years old, adults 19 to 70 years old. Numbers are in percentages

2.3.1 Descriptive Statistics

To study the time use and labour supply behaviour of the spouse and children in response to migration, firstly some characteristics of the sample households and individuals in the study are presented in Table 2.2.

Using the 1997 to 2011 waves from the CHNS data, this paper aims to examine the rural households with at least one child aged 6 to 18 years old and where the mother and children are observed at home. Children and mothers, who appear at least twice in the survey are kept in the study sample. In the analysis sample, rural households are randomly drawn from 146 villages from 9 inland provinces in China. A total of $2,662$ children and $1,787$ women are included in the sample. The children-year observation is 7,226 and the women-year observation is 5,266. There are 1,437 boys and 1,225 girls. The gender ratio reflects the preference for boys as a social norm in rural China, where males are the main source of labour. 1,725 observations are discarded as they are only observed once. These attritors are compared with the sample used in the analysis based on some observables. The t-test shows that the attritors (those observations observed only once) and the non-attritors (those observations observed at least twice) are not systemically different. Thus, this provides some evidence that the discarded observations would not lead to a serious bias in the analysis. Details on the comparison can be found in Table 6.3 in Appendix 6.2.

Within the sample households, the male household head's migration rate is 8.19% during the waves. While 468 out of $2,662$ children (18%) experienced their father's migration status change over the years. One of the main drawbacks of the individual fixed effects model is that a large proportion of data is discarded if the migration status of the male household head does not change. 18% of observations of the children sample saw the father's migration status change and 20.5% of observations of the women sample saw the husband's migration status change. The percentages are relatively large in order to show that there is enough variation in the migration status change for the individual fixed effects model. While the average age of the children is around 12, the average age of the men and women is around 40 years , while migrants themselves are younger.

Table 2.2: Summary statistics of control variables

Variable	mean	sd	N	migration = 0	migration =1	difference
Children age	12.28	3.37	7226	12.28	12.34	0.06
Children education	0.37	0.58	7226	0.38	0.31	-0.07***
Children school enroll	0.88	0.32	7226	0.88	0.92	0.04***
Family size	4.42	1.22	7226	4.40	4.56	0.16***
Male hh head age	39.67	5.87	7226	39.73	39.12	-0.61***
Female hh head age	38.33	5.57	7226	38.36	38.01	-0.35*
Grandmother living	0.19	0.39	7226	0.19	0.21	0.03**
Traffic vehicle	0.84	0.36	7226	0.84	0.86	0.02
Farm machine	0.29	0.45	7226	0.29	0.30	0.01
White goods	1.00	0.99	7226	1.00	1.00	0.00
Information	0.97	0.17	7226	0.97	0.97	0.00
Chore tools	1.82	0.93	7226	1.79	2.05	0.26***
Urbanisation index	52.70	17.18	7226	52.95	50.48	-2.48***
Population density	5.56	1.25	7226	5.56	5.60	0.04
Diversity	4.35	1.00	7226	4.34	4.43	0.08**
Economics	4.55	3.04	7226	4.53	4.76	0.23**
Health	4.66	2.10	7226	4.71	4.21	-0.50***
Housing	5.35	2.37	7226	5.33	5.54	0.21**
Market	4.41	3.74	7226	4.53	3.31	-1.22***
Social service	1.82	2.07	7226	1.79	2.12	0.34***
Transportation	5.19	2.45	7226	5.23	4.86	-0.36***
Education	2.81	1.05	7226	2.84	2.54	-0.30***
Modern markets	3.78	3.09	7226	3.86	3.09	-0.77***
Sanitation	5.03	2.85	7226	5.07	4.70	-0.37***

CHNS, 1997-2011. Children are 6 to 18 years old. Family size is the number of family members living in the household. Male household head is the father of the children, while female household head is the mother of the children. Grandmother living is an indicator of whether the grandmother lives together with the family. Traffic vehicle, farm machine, white goods, information and chore tools are a set of household assets variables. Variables from urbanisation index to sanitation are scores for villages reflecting the social-economic characteristics of the villages constructed by the survey team. The Appendix 6.1 provides detailed explanation of the variables.

The education levels of children, when the father migrates, are lower, but the enrollment rate is significantly higher. The average size of the household is around 4 to 5, and about 20% of the households includes the grandmother (mother of the household head), who is living with the family, while less than 15% of the households have the grandfather (father of the household head) living with them.

This paper focuses on the left behind children and spouse of the migrants. However, as in rural China, it is common for extended family members, such as the parents of the household head, to live together in the household. The extended family members would share some of the household productions, especially when there is a lack of labour. Some research specifically focuses on the labour market behaviour and health situation of the left behind parents of the migrants. One may argue that this study should also compare the time use of the left-behind parents with their counterparts. However, this paper only aims to examine the left behind children and spouse of the migrants. Firstly, this study focuses on the core family members, not the extended family. In rural China, the extended family may indeed help in some activities, such as taking care of infants or doing some household work. Sometimes, other relatives such as uncles/aunts, siblings, and even neighbours within the villages, help each other on the above-mentioned activities. This study aims to focus on the core family member to make both the control and treatment group as clear as possible. Secondly, the left behind parents, compared with the core left behind family members, are the minority in the left behind the group. As statistics from the data show, about 20% of the sample households have either a mother or father of the household head living within the household. In these households, with at least one parent of the household head living with them, about 20% have experienced migration of the male household head. The same statistics apply to the non-migrant household. Thirdly, the analysis in this paper included a dummy variable for the grandmother's living in the household to control for the impact the left behind parents on the time use of the left behind children and spouse. There is a greater tendency for the grandmother to be living in the household than the grandfather. Also, if grandfather lives in the household, in all probability, the grandmother would be staying as well. Finally, the rural elderly residents, who have been allocating their time use for

almost their whole life, would be very unlikely to reallocate. However, their mental and physical health would easily be impacted as they grow weak and lonely. The study of the health and psychology outcomes of left behind parents would be more interesting, as compared to their time use outcomes.

The household assets measures are traffic vehicles, farm machines, white goods, information and chore tools. Traffic vehicles is an indicator showing whether the family owns a tricycle, bicycle or motorcycle; farm machines is an indicator for owning a tractor, water pump, irrigation machines or thresher. White goods are the number of machines the family owns, be it a washing machine, fridge or air conditioner. Information is the indicator for ownership of computer, telephone, mobile phone and TV. Chore tools are the number of the household appliances the family owns - fan, microwave oven, cook-pot and others. The statistics in Table 2.2 indicate that the migrant and non-migrant households have no significant difference in assets, except the chore tools. On an average, the migrant households own 0.3 more pieces of chore tools.

Some statistics on the outcomes of this study, labour participation and time use for sample children and spouse are presented in Table 2.3. The upper section of the table shows the outcomes of children and compares the outcomes with the father's migration status. The lower part of the table shows the outcomes of the women and compares the outcomes by husband's migration status. The outcomes of interest are the participation and hours per day the children and spouse spend on wage labour, housework and agriculture. The outcome choices are not mutually exclusive, as individuals can choose to work in one or more of the three activities, or none of them. Family members have the overall control of the agriculture product. They can consume the agriculture product by themselves or trade it in the market. Those who are employed and get paid are wage labourers in the rural area (defined as local villages, towns, or suburban villages; they may be employed permanently or temporarily, part-time or full-time), including those who are employed and paid for doing agriculture work for the state, collective or privately owned farms. Participation in wage labour may be constrained by the local employment opportunity, while participation in housework or agriculture is not constrained.

Table 2.3: Summary statistics of outcomes by migration status

Dependent variables	mean	sd	N	migration = 0	migration =1	difference
Children				N= 6,483	N = 743	
Agriculture par	0.06	0.23	7,226	0.06	0.06	0.00
Agriculture hours	0.21	1.06	7,226	0.21	0.18	-0.03
Housework par	0.22	0.41	7,226	0.21	0.28	0.07***
Housework hours	0.18	0.58	7,226	0.17	0.23	0.05***
Wagelabour par	0.02	0.14	7,226	0.02	0.01	-0.01
Wagelabour hour	0.14	1.05	7,226	0.14	0.12	-0.02
Women				N = 4718	N = 548	
Agriculture par	0.59	0.49	5,266	0.58	0.69	0.11***
Agriculture hours	4.11	4.57	5,266	4.01	4.39	0.31*
Housework par	0.93	0.25	5,266	0.93	0.95	0.02*
Housework hours	2.79	1.97	5,266	2.78	2.88	0.10
Wagelabour par	0.47	0.50	5,266	0.46	0.59	0.13***
Wagelabour hour	3.22	3.79	5,266	3.18	3.56	0.38**

CHNS 1997-2011. Children are 6 to 18 years old. N is number of observations, mean is the mean value and sd is the standard deviation. Par and hours are participation in and time use in the three types of activities. Migration is father's migration status for children outcomes and husband migration status for women outcomes.

6% of children participate in agriculture and spend about 0.2 hours per day in agriculture. The time use in agriculture of children in a migrant family and non-migrant family are not significantly different. Rural women are extensively engaged in agriculture. On an average, about 60% of rural women participate in agriculture and they spend about 4 hours per day on these activities. The spouse of migrants significantly participate more in agriculture and spends 0.3 hours more per day.

Participation in housework for children is around 20% during the period observed. Children in migrant families participate more in housework and spend more time in housework, as compared to the children in non-migrant families. The average participation in housework for women is 97% and the average hour is about 3 hours per day. There is no significant difference between the spouses of migrants and those of non-

migrants on the time use in-house work.

Wage labour participation for children is about 2% on average. However, when child migration is considered as wage labour (in destinations of migration), wage labour participation amounts to about 5% to 15% of all rural children aged 6 to 18 years old. Interestingly, one can also observe the tendency for children to migrate than to participate in the local labour market (wage labour), and the growth of the migration rate is high before the education reform during the period 2006 to 2008. The wage labour participation for women is about 47% and the average time per day is about 3 hours. The spouses of migrants are involved more in the wage labour.

It is interesting to note that the children in migrant families participate more in housework, while the spouses of migrants participate more in income-generating activities (agriculture and wage labour). However, the statistics are unable to specify the exact impact of the migration, as many other heterogeneities contribute to the variations in the outcomes.

2.4 Empirical Strategy

In order to study the impact of migration on the time use and labour supply of the left behind spouse and children, consider the following regression model:

$$Y_{i,t} = \alpha + \beta Mig_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \qquad (2.1)$$

where the dependent variable $Y_{i,t}$ is the outcome of interest of left behind individual i at time t. This linear model can model the following 4 scenarios: (i) The binary outcome of the spouse's participation decision, a dummy variable indicating whether or not the spouse participates in each of the three types of work (household work, wage labour and agriculture); (ii) The binary outcome of children's participation decision, a variable indicating whether or not the children participate; (iii) How many hours per day the spouse spends on each activity; (iv) How many hours per day the children spend on each activity.

The explanatory variable $Mig_{i,t}$ is the current migration status of the father of i in the children's sample and the current migration status of the husband of i in the women's sample. It equals 1 if the male household head is a current migrant and 0 if not. $X_{i,t}$ is a set of the characteristics of individuals, households and villages. The list of control variables is presented and explained in Table 2.2. ε includes all the unobservable factors that may affect the outcomes stated above. β and γ are the parameters. Due to the endogeneity of migration because of time-invariant or varying omitted variables and simultaneity, there is a correlation between the migration variable and the error term. This paper first presents the pooled OLS model, and the panel fixed effects model that deals with the time-invariant source of endogeneity, and then propose an instrumental variable to deal with the other sources of endogeneity.

First, the pooled OLS and linear panel fixed effects model are implemented as benchmark models. Table 2.4 presents the pooled OLS and linear fixed effects regression results. The first column indicates whether the model and outcomes are for the left behind children or spouse. Last three columns refer to participation or time spent on the three types of activities. The coefficients and standard errors of $Mig_{i,t}$ are presented in the table. The estimation results of pooled OLS and linear panel fixed effects models show that overall, migration has no significant effect on the time use of the left behind children. The exceptions are the increase in the housework participation of the left behind children and the agriculture participation of the left behind spouse. This is plausible, as the left behind women compensate for the migrants on the family farm, while the left behind children compensates for their mother in the housework. However, the insignificant and significant coefficients are mostly in small magnitude, about zero. The benchmark results show the weak influence of migration on the time use of the left behind children and spouse.

Table 2.4: Bench mark models: pooled OLS and panel fixed effects models

	agriculture		housework		wage labor	
	participation	hours	participation	hours	participation	hours
Children outcome Pooled OLS model	-0.017	-0.042	0.008	-0.020	-0.006	-0.020
	(0.009)*	(0.037)	(0.017)	(0.023)	(0.004)	(0.040)
Children outcome Panel Fixed effects	-0.014	0.006	0.047	0.010	-0.006	-0.015
	(0.015)	(0.048)	(0.023)**	(0.032)	(0.006)	(0.059)
Spouse outcome Pooled OLS model	0.051	0.043	0.008	-0.003	0.025	-0.139
	(0.019)***	(0.035)	(0.010)	(0.030)	(0.020)	(0.156)
Spouse outcome Panel Fixed effects	0.049	0.078	0.013	-0.059	0.052	0.020
	(0.023)**	(0.043)	(0.014)	(0.040)	(0.028)*	(0.196)

CHNS, 1997-2011, child 6 to 18 years old. $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses.

Variable $Mig_{i,t}$ is a choice variable, and in the case of rural-urban migration in China, it is mainly due to the economic reasons. Questions arise as the explanatory variable $Mig_{i,t}$ is correlated with the error term ε. Thus the OLS estimation can be biased. Migration is a self-selection decision. One potential source of endogeneity is some unobserved individual heterogeneities subsumed in the error term, which may influence both the migration decision and the outcome of the left behind. For instance, some children may be more interested and able at school, thus they may intend to spend more time learning, whilst participating less in housework, farming or other activities. The father may have a stronger incentive to migrate, in order to generate more income for the education investment. Another situation may be that, for some rural women, their individual characteristics (such as skills and language ability) may force them to keep working on the farm or do housework, and participate less in the labour market or running the family business. In the meantime, the husband may need to migrate to generate more income for the family.

The panel nature of the data provides a method to control for this type of endogeneity. The individual fixed effects method is applied in order to eliminate the time invariant sources of endogeneity. By using the individual fixed effects methods, the potential endogeneity that may attribute to the observed or unobserved time invariant variables, such as individual gender or ability, are controlled for. Besides, time invariant household or village characteristics, whether observable or unobservable, are controlled at the mean

time. By including the time fixed effects and writing the error term as $\varepsilon_{i,t} = \eta_i + \nu_{i,t}$, the following equation is derived:

$$Y_{i,t} = \alpha + \beta Mig_{i,t} + \gamma X_{i,t} + \eta_i + \tau_t + \nu_{i,t} \qquad (2.2)$$

where η_i is the individual level factors, both observable and unobservable, for the children and the spouse. The fixed effects model allows for correlation between the time-invariant variables with the endogenous variable. Furthermore, the time fixed effects τ_t is included.

The endogeneity of migration may also come from other time variant unobserved factors or pull and push factors that can not be captured by η_i. Thus, the fixed effects model may also be biased. These time-variant factors lead to a correlation between the male household head's migration and $\nu_{i,t}$. If the motivation for migrating is the vulnerability of the household, for instance, an unpredictable shock in agricultural production (which may result from floods, droughts, pests or the rainfall), it may push the male household head to migrate, the spouse may participate more in wage labour and the children may need to spend more time on household chores and wage labour. The complexity of analysing the impact of migration is rooted in the fact that families make decisions about migration simultaneously with decisions on labour supply, education investment, child labour, household farming and so on.

To control for the time variant source of endogeneity, an instrumental variable (IV) is proposed. To be a good candidate, the IV needs to show strong correlation with $Mig_{i,t}$ and no correlation with $Y_{i,t}$. In other words, the IV should affect the outcomes only through its impact on migration. The IV applied in this paper is based on the labour market conditions in the coastal cities of China, where migrants most likely migrate. The linear model above is estimated using an IV with individual and time fixed effects. The IV is the wage gap between the relatively developed coastal cities and the local urban areas around the migration-sending villages and it is clear that the orthogonality condition is likely to be satisfied. The wage gap is measured as follows:

$\log W_{s,U,t-1}$-$\log W_{e_i,u,t}$

Where $W_{s,U,t-1}$ refers to the lagged one-year average annual wage in five popular sectors s in popular destinations in coastal areas U, where integer $s \in [1, 5]$ is the sector domain and U is the set of coastal cities, {Beijing & Tianjin, Shanghai & Zhejiang, Guangdong}, t is the survey year, and $t-1$ is the time point by the end of t-1. $W_{e,u,t}$ is the average annual wage in local urban areas u sorted by the education level of individual e_i at time t, where u refers to the local urban areas located within the same region as the migration-sending villages, e is the education categories, 0 for primary level or lower, 1 for middle school, 2 for high school and above. Thus, the instrumental variable has several dimension of variation by education level of individuals, geographical distribution of local villages, time dimension and industry dimension of the sectors. There are sufficient variations in the instrumental variable adopted for the fixed effects model, as the overall variance of the IV is 0.34 and the within variance is 0.21.

The popular destinations of coastal areas where the migrants are most likely to migrate are Beijing, Tianjin, Shanghai, Zhejiang Province and Guangdong Province. The five sectors that comprise the industries, where migrants are most likely get employed are manufacture, construction, transportation, retail, food and accommodation. Local urban areas are suburban, towns or small/medium cities in the same province as the migrant sending villages identified in the survey. The wages are real wages in 2011 currency rate.

It has been widely accepted that the wage gap and income inequality between rural and urban areas, and across coastal and inland parts of China, is an important factor of rural to urban migration in China. Zhu (2002) finds evidence to support the hypothesis that the "urban to rural income gap is important and significantly influences migration decisions". A recent study by Chan (2013), like many papers reviewed in it, states that "labour migration flows are closely linked to significant disparities in wages between urban and rural sectors and between regions in China".

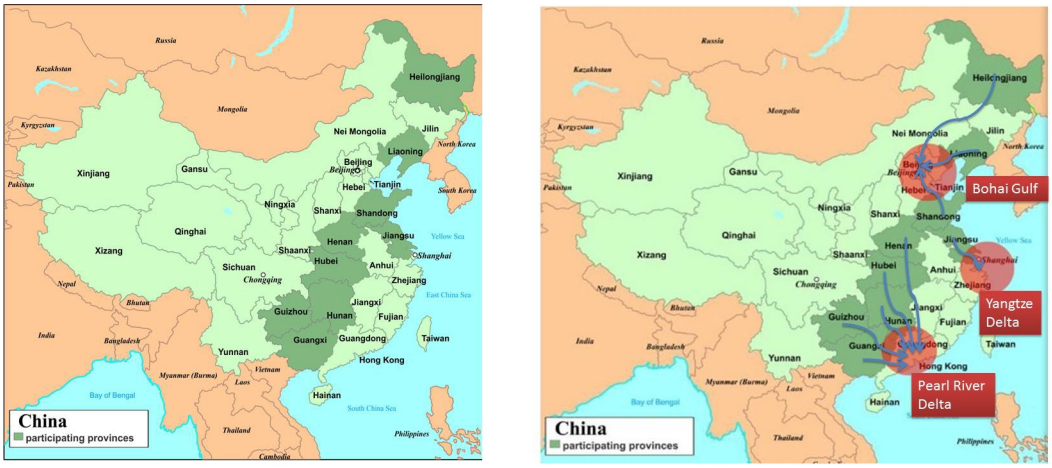
Migration is self-selected and a choice variable. Potential migrants form an expectation about the income if they stay in their home villages, as well as the income if they migrate based on the wages observed at the most likely destination (W_{t-1}). The papers

use the lagged one year wage in the destination for the following reasons. First, potential migrants can observe the wage of $t-1$ which is revealed, while it is hard for them to obtain information on current wages at the moment of migration decision; second, the time difference in the two wages adds more variation in the IV and makes the IV more exogenous to the current outcomes of the left behind family members. It is straightforward to argue that the instrumental variable is strongly correlated with the current migration status observed at time t for individual i. The wage gap between local and coastal cities does not impact the outcomes of the left behind directly in rural areas and the exclusion condition holds. On the orthogonality of the instrumental variable, this paper uses the geographical distribution of surveyed provinces, the segregation of the rural and urban labour markets, in order to address the exogeneity of the instrumental variable. Details are discussed in the following section.

2.4.1 Instrumental Variables

The construction of the instrument begins from a number of facts about internal migration in China. As the International Labour Organisation (ILO) states: "most internal migrants in China are rural to urban migrants and the migration flows basically from the interior to coastal areas".

Chan (2013) concludes that "since the early 1990s, the inter-provincial migration has increased rapidly and gained popularity over time, and the migrants have a clear tendency to concentrate in Guangdong (the Pearl River Delta) and Shanghai (the Yangtze Delta)". Poncet and Zhu (2005) find that "31% of the rural migrants moved towards coastal areas during 1985-1990, and the proportion increased to 47% over the period 1990-1995 and 62% over the period 1995-2000". Fei (2007) concludes that "rural to urban migration is increasingly concentrated in urban areas and coastal regions with 65% of rural migrants moving to coastal areas and 84% of inter-provincial migrants moving from central to eastern regions in 2000". Poncet and Zhu (2005) indicates that most migrants favour three destinations, the Pearl River Delta (Guangdong), the Yangtze Delta centred on Shanghai, and the Gulf of Bohai centred on Peking and Tianjin.



Based on maps from CHNS official page, Carolina Population Center

Figure 2.1: Survey provinces and the rural to urban migration flow

Based on the migration flow and trends in China, Figure 2.1 shows how this paper matches the survey provinces with the three coastal regions, the Pearl River Delta in the south coast represented by Guangdong Province; the Yangtze Delta in the east coast represented by Shanghai and Zhejiang Province and the Gulf of Bohai in the north coast represented by Peking and Tianjin.

The migrants from Heilongjiang, Shangdong and Liaoning are most likely to migrate to the northern coastal region, represented by Peking and Tianjin. Migrants from Jiangsu Province mostly migrate to the east, where Shanghai and Zhejiang are economic centres. Migrants from Henan, Hubei, Guizhou, Hunan and Guangxi are most likely to migrate to the south to Guangdong. These assumptions of migration patterns are supported by the statistics of the 2000 and 2010 census in China, presented in Table 2.5.

From the 2000 and 2010 census from the National Bureau of Statistics in China, this paper constructs migration flows from the survey provinces to the coastal regions. Table 2.5 shows the percentage of rural migrants from the survey provinces who migrated to a destination (the matched coastal region). The "numbers" presented are the total migrants sending from the surveyed provinces. It is important to point out that the

proportion of migrants moving to the matched destinations (in percentage in the table) is the largest among the percentages of migrants going to any other destinations. One exceptional case is Heilongjiang, where about 30% of the migrants move to other cities in Gulf of Bohai, while 15% (the largest proportion) migrate to Beijing $\&$ Tianjing in 2000 and 28% in 2010. This does not mean it is not corresponding, as more than 50% of migrants from Heilongjiang move to the Gulf of Bohai. The wage gap between local urban areas in Heilongjiang and the Gulf of Bohai can be represented by the gap between the local urban areas and the representative cities in the Gulf of Bohai, which are Beijing and Tianjin.

Migrants from Guizhou show a similar preference for the east and south. In 2000, Guangdong was the most popular destination, while in 2010 it was Shanghai $\&$ Zhejiang; 36.63% went to Guangdong in 2000 and in 2010, the majority 45.07% went to Shanghai and Zhejiang in the east.

Table 2.5: The migration flow from survey provinces to coastal regions

		2000		2010	
Home	Destination	number	percentage	number	percentage
Liaoning	Beijing & Tianjin	15,210	14.62%	41,902	28.32%
Heilongjiang	Beijing & Tianjin	47,570	10.62%	137,684	19.73%
Jiangsu	Shanghai & Zhejiang	94,272	50.15%	211,819	68.07%
Shangdong	Beijing & Tianjin	66,596	26.61%	207,736	31.51%
Henan	Guangdong	203,544	35.84%	435,217	35.38%
Hubei	Guangdong	172,065	54.55%	445,602	43.15%
Hunan	Guangdong	278,627	79.20%	595,969	67.07%
Guangxi	Guangdong	166,016	90.59%	358,628	87.63%
Guizhou	Guangdong	112,604	36.63%	345,906	23.69%
Guizhou	Shanghai & Zhejiang	112,604	20.79%	345,906	45.07%
total		1,156,504		2,780,463	

Source: 2000 and 2010 Population Census of China, National Bureau of Statistics

This reflects migrants' increasing interests in the east, which is a recent tendency. As

the data in this paper covers the duration from 1997 to 2011, it is reasonable to assume that most of the migrants from Guizhou went to Guangdong during the survey time. It is interesting to note that almost all of the rural migrants moved to the developed coastal region, close to their home provinces, except for those from Henan and part of Hubei province in the central area. This is consistent with migration literature that distance and the moving costs are both important determinants of migration flow. The exceptional case of Henan and Hubei, where most of the rural labourers migrate to Guangdong, may root in the fact that south to north train travel was much more convenient than west to east train travel in China during that time period. Wuhan, the capital city of Hubei province and Zhenzhou, the capital city of Henan, are located on one of the most important railway lines in China, the Jingguang line, which directly links the rural labourers in these two agricultural provinces to the Pearl River Delta. Besides , Guangdong is where the "Open and Reform" policy started and is more migrant-friendly, as compared to the other destinations, like Peking.

The wages for the destination cities are the average real wages of five popular sectors in the five cities/provinces clustered in the three most likely and representative destinations of the coastal regions. Migrants not only have the tendency to cluster in regions geographically but also in the five industrial sectors of manufacturing, construction, transportation, retail, food and accommodation. According to the China Labour Bulletin, (as shown in Figure 2.2), during 2008 and 2012, the sector-wise distribution of rural migrants is almost stable, with over 80% of the rural migrants working in manufacturing, construction, transportation, retail, food and accommodation.

Figure 2.3 shows the average real wages in the five sectors at the three migration destinations. The wages are in 2011 currency values. The wages in the coastal regions increase steadily over the survey years. One may argue that the living expenditure in the coastal urban regions is higher. This is true, however, considering the strong preference for saving over consumption, even though the average living expenditure in coastal cities increases, rural migrants always save and remit more.

As the fixed effects instrumental variable (FEIV) estimation is applied in this paper,

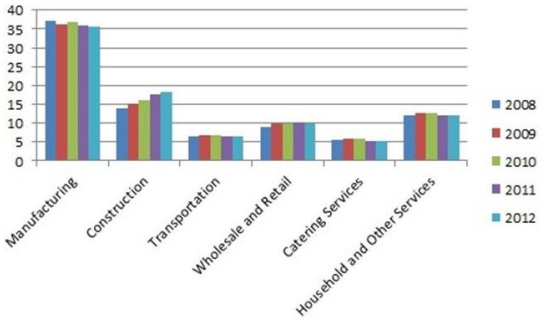


Figure 2.2: Popular sector for migrants

Figure 2.3: Average wage in five sectors in destinations

sufficient variations in the instrument are required. The IV varies for individuals and changes over time and across regions; plus, when considering the difference of the real wages, it also adds variation to the instrument. Figure 2.4 shows the log wages in local urban areas, which is the average annual wage in local urban areas, sorted by the education level of the individual. Figure 2.5 shows the variation in the IV, where the differences of the log wages are above zero. $\log W_U$-$\log W_u$=$\log(\frac{W_U}{W_u}) > 0$ and most wage gaps are over 0.5, which means the real wage in coastal cities is more than 3 times the real wage in local cities. The wage inequality in inland and coastal cities increases over the year. This also explains the rapid increase in migration rates.

The relationship between the instrument and the father's migration is straightforward; higher income is the main motivation for migration. The instrument explains the migration of rural labourers to urban areas and explains why the developed coastal cities are more attractive. The correlation can be verified by the first stage regression results presented in the main results, and the F statistics on the excluded instrument indicate the strength of the IV applied in this paper.

It is important to address the orthogonality and discuss cases that may violate the exclusion restriction. The exogeneity of the instrument is based on the arguments the segregated rural and urban labour market in China; geographical distribution of survey

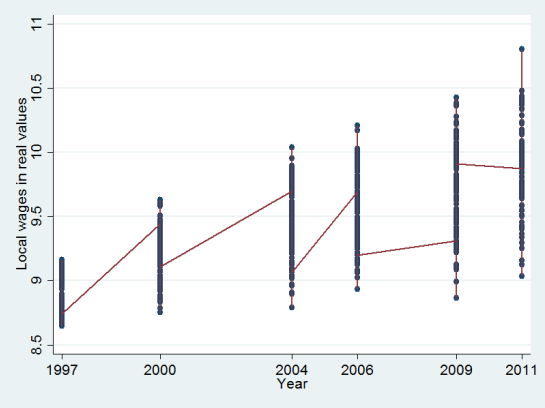


Figure 2.4: Log wages in local urban areas

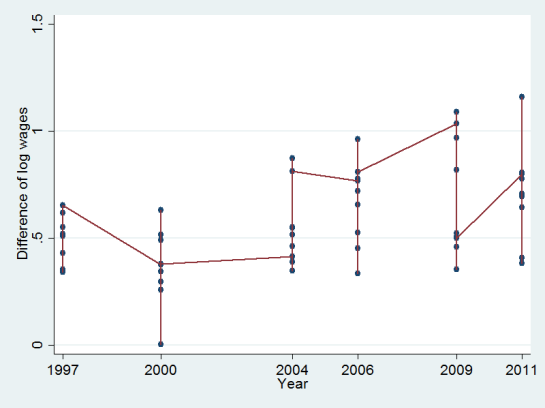


Figure 2.5: Wage gap variation

provinces and migration flow; the lagged value of the wage gap; and the real wage sorted by the education level of the individual.

Urban wages are not related to the left behind children and spouse's outcome in the migration-sending villages due to certain policy constraints. In the short term, the rural and urban labour markets are strictly segregated and linked by rural to urban migrants. Urban labour market conditions do not directly affect the left behind children and spouse's outcomes.

From Figure 2.1, it is clear that rural migrants move from survey provinces to the coastal regions. The geographical distribution of the survey provinces that are inland provinces, and migration destinations that are coastal regions, adds to the exogeneity of the instrument. The wage difference between the coastal and local urban areas is orthogonal to the error term in the model. The lagged value of the wage gap makes the instrumental variable more exogenous with the current time use behaviour of the left behind. The wages in the local urban areas are sorted by the education level; this means the individuals with different education levels in the same village have different values for the instrumental variable.

2.4.2 Threats to Identification

One challenge to the exogeneity of the instrument is the fact that local cities, coastal cities, and the migration-sending villages face the same economic environment in China. The Chinese business cycle affects the coastal cities, local cities and the villages together, especially when the economy is more centrally controlled. Some unobserved factors, due to macroeconomic fluctuations and the institutional factors, may impact the wage levels in coastal and inland cities, as well as the left behind children and spouses' time use and labour supply. However, considering the difference between wages in coastal cities and those in inland cities eliminates the effect of the business cycle.

The other threat to identification is the fact that the instrument is also correlated with the migration probability of the children and spouse. Although most of the migrants are male heads of households and they have the highest incidence of migration, the children and spouse left behind may also migrate. People in migrant families with migration networks may have a higher probability to migrate. Thus, their time use and labour supply may be different. However, even though the left behind may have a higher probability to migrate if the male household head migrated, their current left behind status is observed in the data and, even before migration, they still distribute their time into types of work in the short term as usual. This study looks at the short term adjustment of time use of the left behind, but not the long term effect of migration.

To sum up, the instrumental variable does not impact the outcomes directly and is not correlated with the error term, while the orthogonality and exclusion condition of the instrumental variable are argued to be satisfied. After dealing with the time invariant and the time variant source of endogeneity of migration, the two-stage regression with the IV and FE (FEIV) is implemented. Let $Z_{i,t}$ be the instrumental variable, first stage regression in the fixed effects instrumental variable would be the following:

$$Mig_{i,t} = \pi Z_{i,t} + \theta X_{i,t} + \phi_i + \tau_t + \epsilon_{i,t} \qquad (2.3)$$

$Z_{i,t}$ is the wage gap constructed to instrument migration, π, θ are the parameters. The main results are presented and interpreted in section 2.5.

2.5 FEIV Main Result

The two-stage least squares regression with 1, an instrumental variable for the single endogenous variable; 2, the individual fixed effects; 3, time fixed effects, is implemented with the STATA command Correia (2014) "reghdfe", which is a user-written package. In this paper, the dependent variables are actually binary variables and censored outcomes. The linear regression is applied because the non-linear models can not accommodate the combined instrument variable and fixed effects estimation. Actually, the outcomes in the study are a discrete choice or censored data, and some may prefer the Probit/Logit model for the discrete choice or Tobit model for the censored outcome. The nonlinear models like Probit/Logit and Tobit produce conditional expectation functions that respect limited dependent variable (Angrist and Pischke (2008)). However, the use of the individual fixed effects in the models prohibits the use of limited dependent variable maximum likelihood methods (Antman (2011a)). Furthermore, there are some advantages for using linear models for the nonlinear outcomes. First, the linear models allow this paper to apply the individual fixed effects and the instrumental variable approach to deal with the endogeneity issues. Second, it is easy to estimate and interpret in terms of marginal effects. Third, robust clustered standard errors in this study would easily deal with the heteroskedastic problem in the linear regressions, while robust cluster standard errors may be difficult to implement in some nonlinear models. Fourth, comparing to choose any arbitrary nonlinear model without clear mind about the conditional expectation function forms, linear models can be a preserved but safe choice. Besides these, some may say why not the Heckman model but the IV approach. The Heckman model needs a strong assumption about the joint normal distribution, which is doubted by more and more economists, while the IV approach is more robust to distributional misspecification. One may also ask why not the first difference models, but the fixed effects models? The answer is when time span is 2, the two types of models are the same. However, when time span is long, the fixed effects models are more efficient, Wooldridge (2010). Also if the time span is long, there may be a problem with first difference method regarding the serial correlation of the error terms. Above all, the linear

model with fixed effects and the instrumental variable is applied in this analysis.

Table 2.6 and Table 2.7 present the main results on how the father/husband's current migration status affects the time use and labour supply of the left behind children and spouse. The dependent variables are the time use with regard to three activities, wage labour, housework and agriculture. In each of the equations, individual, household and village characteristics (for details about the explanation and summary statistics of these controls, please check Table 2.2 and Tables 6.1 and 6.2 in Appendix 6.1) and time dummy are included.

The first stage regression is the regression results of equation 2.3. The positive and significant coefficients π of the instrumental variable $Z_{i,t}$ in the first stage regressions (0.06 in Table 2.6 for children outcomes and 0.04 in Table 2.7 for women's outcomes) shows the strong relationship between the instrument (wage difference) and the migration status. This strong relationship suggests the instrumental variable, as addressed in section 4, is an important determinant of the migration decision.

In these estimations, there is only one instrumental variable for one endogenous variable; the equations are exactly identified, and the Sargan statistics are zeros. It may be useful to apply multiple IV for over-identification test but feasible IVs are not easily available. The F-statistics of the excluded instrument in the first stage are large. The F statistics are 12.50 for regressions in Table 2.6 for children's outcomes, which is larger than the rule of thumb value of 10, hence, there is no need to worry about the weak instrument issue. The F statistics are around 5 in the regressions in Table 2.7 for the spouse's outcomes. However, the p values are smaller than 0.05, and the statistics are significant. The R-squares of the regressions are relatively large. Considering these aspects, the weak instrumental variable is not an issue in the women's regression either. The wage gap does not seem to be a weak instrumental variable for the father/husband's migration status. As the first stage statistics are under 10, the results for the left behind spouse need some caution while interpreting.

As the instrumental variable methods are applied in these regressions, the test of endogeneity of the endogenous variable has been implemented and reported in the re-

gression tables. The test implemented is the "Durbin-Wu-Hausman test of endogeneity" (Durbin (1954), Hausman (1978),Wu (1974)), which reports the Durbin-Wu-Hausman Chi Square test. The results show that the migration of the male household head should be treated as an endogenous variable.

The second stage results in Tables 2.6 and 2.7 present detailed results of the main exercise. Overall, the migration of the household head has no significant impact on the time use of the children and women in all three types of activities, neither in participation nor the hours spent, except the time use in children's wage labour. The results are actually not surprising, as the income effect and the substitution effect of migration on the left behind work in opposite directions. The income effect cancels out the substitution effect of migration, thus the net effect of migration on the left behind is found to be non-significant. The non-significant results do not mean that migration doesnt affect on the left behind, but indicates that the income and the substitution effects of migration offset each other.

As presented in column "agriculture" of Table 2.6 and 2.7, the coefficients of children and women's participation and time spent on agriculture are positive (0.53 & 2.33 for children's agriculture outcome; 0.75 & 1.88 for women's agriculture outcome). However, these coefficients are not statistically significant.

Before interpreting the left behind's time use in agriculture, it is vital to point that farm, which is distributed to a household's per capita, is an important endowment of the rural households. The limited size of the land, the abundance of the family labours, and the low marginal productivity of agriculture are push factors of migration. However, the land will always be treated as a household asset, an input for production and a backup plan for the family averse to risk. Traditionally, when the male household head migrates, the spouse always needs to stay on the land for autarky or generating some income. However, as cultivable land (or ponds, lakes, etc.) is distributed among rural households by the government, they can choose to farm or not. The right to use the land does not imply an obligation. In the 1990s, large numbers of farmers abandoned their land, regardless of their choice to migrate or stay, because of the agricultural tax. To

ensure food security and social stability, the Chinese government started reducing the agricultural tax since 2000 and finally cancelled it in 2005. [3]

Table 2.6: FEIV estimation: the impact of father's migration on children's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.060***	0.060***	0.060***	0.060***	0.060***	0.060***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
First stage F-statistics	12.50	12.50	12.50	12.50	12.50	12.50
Prob > F	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Second Stage						
Father migration status	0.532*	2.329*	-0.527	-1.189	0.428**	3.095**
	(0.285)	(1.215)	(0.445)	(0.727)	(0.194)	(1.499)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	19.90***	19.90***	19.90***	19.90***	19.90***	19.90***
Weak identification test	12.50***	12.50***	12.50***	12.50***	12.50***	12.50***
Weak instrument robust inference	8.66***	10.34***	2.74*	7.20***	14.31***	11.68***
Durbin-Wu-Hausman test	5.09**	5.24**	6.04**	3.74*	8.13***	6.41**
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	2,667	2,667	2,667	2,667	2,667	2,667
R square	0.288	0.314	0.418	0.280	0.054	0.098
N	7,226	7,226	7,226	7,226	7,226	7,226

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

Looking at the results in the "agriculture" column of Table 2.6 and 2.7, the time use for agriculture of the left behind children and women is not different from that of children and spouses in non-migrant households. The impact of migration on the agricultural activities of the left behind is not significant, suggesting that the income and substitution effects of migration cancel out and the net effect of migration is, hence, not significant.

[3]The Official site of the Central Government of PRC: $http://www.gov.cn/test/2006-03/06/content\_219801.htm$

Table 2.7: FEIV estimation: the impact of husband's migration on spouse's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.041**	0.041**	0.041**	0.041**	0.041**	0.041**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
First stage F-statistics	4.98	4.98	4.98	4.98	4.98	4.98
Prob > F	0.026**	0.026**	0.026**	0.026**	0.026**	0.026**
Second Stage						
Husband migration status	0.748	1.876	0.285	2.129	-0.201	2.125
	(0.674)	(1.449)	(0.416)	(1.403)	(0.692)	(5.280)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	7.64***	7.64***	7.64***	7.64***	7.64***	7.64***
Weak identification test	4.98**	4.98**	4.98**	4.98**	4.98**	4.98**
Weak instrument robust inference	2.45	3.94**	0.89	6.67***	0.13	0.26
Durbin-Wu-Hausman test	4.52**	4.21**	5.37**	3.68*	4.96**	5.64**
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	1,792	1,792	1,792	1,792	1,792	1,792
R square	0.641	0.523	0.402	0.013	0.619	0.614
N	5,266	5,266	5,266	5,266	5,266	5,266

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

The substitution effect of migration would lead the left behind children and the spouse to work more on agriculture when the main labour of the household is absent. A previous study, Mu and Van de Walle (2011)) concludes that left behind women are always expected to do the farm work. However, due to the existence of the income effect, the left behind spouse and children may not work so much on the farm, as the migration of the male household head brings in more income for the family. Remittances can be used to rent agricultural labour or machinery in order to compensate for the lost labour caused by migration and to reduce the left behind children and women's time allocation for agriculture. Considering the two opposing effects on migration, it is simple to conclude the overall non-significant effect of migration.

There is no evidence that the husband's migration decreases the wage labour of the

spouse. The left behind spouse's labour supply to wage labour is not statistically different from the spouse of a non-migrant male household head. Similar to the agriculture behaviour, the migration of the husband has two effects on the left behind spouse; the income effect, which will decrease the labour force participation, and the substitution effect, which would make the left behind spouse participate more and spend more time on wage labour. The result indicates that the income and substitution effects co-exist and offset each other. Thus, the labour supply to wage labour of the left behind spouse is not statistically different from the labour market response of the spouse in non-migrant households.

The columns "housework" in Table 2.6 and 2.7 show the children and women's time allocation for housework. The coefficients are negative for children's outcome (-0.53 $\&$ -1.19) and positive for women's outcome (0.29 $\&$ 2.13). The non-significant results suggest that the substitution effect cancels out with the income effect. The substitution effect would increase the time use in housework. On the other hand, the income effect is also expected to impact the behaviour of the left behind. For instance, more income in the family may lead to a higher level of household production technology (with more white goods, chore tools, etc.).

Furthermore, the impact of the father's migration on children's wage labour is positive. Children of migrant family increase around 43% in the participation incidence, and spend about 3 more hours per day on wage labour as compared to the children in the non-migrant family. The husband's migration has no significant impact on the wage labour of the left behind women (the coefficients are -0.20 $\&$ 2.13, respectively). However, the children's time use in wage labour increases with the father's migration. The income effect results from remittances sent by migrants and decreases their children's wage labour, while the substitution effect of migration may lead the left behind children to work more in wage labour. The migration of male household heads in the rural labour market leads to a decrease in labour supply in the local labour market, and the children may be considered as compensation by the local factories for jobs requiring few skills or experience. Antman (2011a) concludes that there is an increase in child labour in response to the father's international migration, and this is due to the family

undergoing financial hardship in the immediate aftermath. This interpretation seems to contradict the income effect of migration, however, it can be true that the cost of migration and the delay in receiving remittances would make the left behind children take up temporary paid jobs. It can be the case that the migration of the male household head may cause financial difficulties for a short period of time immediately after migration for the household. The migration of the male household head also decreases the labour supply in rural labour markets. Thus, the left behind children may do more wage labour. For the children, time used in wage labour, the income effect and substitution effect still work together; however, there may be a delay in the timing of the income effect. Maybe, when the remittance is received a short period after migration, the left behind children would stop or decrease work in wage labour, thus the income effect trades off with the substitution effect.

The finding of an increase in children's time use in wage labour is different from findings of Chang et al. (2011) which using the same data as this paper. They find rural to urban migration increases the time spent on the farm and domestic work for the left behind children but do not study the time use in wage labour or off-farm work for the left behind children. Moreover, although using the same data with this study, Chang et al. (2011) use the pooled cross-sectional data instead of panel property of the data this paper exploits, and they use migration rates in the village as instrumental variables, which are supposed to impact the labour supply of the left behind, thus this creates certain doubt on their methodology.

Overall, the impact of migration on the time use of the left behind children and spouse is not significant. This suggests that the income effect and the substitution effect of migration trade off against each other. Thus, the net effect of migration is not significant. The following parts would support this main finding.

2.5.1 Household Income

Some may argue that the father's migration is closely related to household income. Poorer farmers may have a stronger incentive to migrate for higher income. It is also acknowledged that household income relates closely to the family member's time use and labour supply. In this study, the household income may be one factor that leads to the selection of migration. The household income, simultaneously, affects the migration decision and the outcomes of the left behind. In this exercise, the household income per capita is added as a control, to determine how the impact of migration on the time use and labour supply of the left behinds changes.

It's noticed that including endogenous control variable in the model is not a good choice. Adding endogenous control variable may lead to biased and inconsistent estimators, but simply omitting them would lead to omitted variable bias. In the model of this paper, household income is actually an important determinant of time use and labour supply of left behind, but it is also endogenous (correlated with the error term) and it directly impacts the migration decision of the male household head. When control variable is endogenous, it's very likely that the bias of including it is greater than omitting it, Angrist and Pischke (2008). Thus the main analysis of this paper excludes the endogenous control. However, this analysis would like to show the results when the endogenous control is excluded or included. The results are consistent in both situations, suggesting that the instrumental variable is independent of what the model controls for.

Tables 2.8 and 2.9 show the results after the household income is added as an explanatory variable. The household income is annual per capita income in 2011 currency value. It is clear that, when the household income is added as a control, the fixed effects estimation with the proposed instrument produces almost the same results as the main results presented in Tables 2.6 and 2.7. The main finding of this paper is robust to the household per capita income. The time use and labour supply of the left behind children and spouse are not significantly affected by migration, suggesting that the income and substitution effects work together in opposite directions and trade off each other, except that the left behind children would increase the wage labour supply.

Table 2.8: FEIV with income variable: the impact of father's migration on children's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.058***	0.058***	0.058***	0.058***	0.058***	0.058***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
First stage F-statistics	12.00	12.00	12.00	12.00	12.00	12.00
Prob > F	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
Second Stage						
Father migration status	0.542*	2.369*	-0.550	-1.232	0.445**	3.222**
	(0.294)	(1.253)	(0.458)	(0.750)	(0.201)	(1.554)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	19.12***	19.12***	19.12***	19.12***	19.12***	19.12***
Weak identification test	12.00***	12.00***	12.00***	12.00***	12.00***	12.00***
Weak instrument robust inference	8.61***	10.29***	2.85*	7.46***	14.75***	12.03***
Durbin-Wu-Hausman test	5.04**	5.18**	5.09**	3.80*	8.31***	6.56**
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	2,667	2,667	2,667	2,667	2,667	2,667
R square	0.280	0.308	0.412	0.270	0.027	0.074
N	7,226	7,226	7,226	7,226	7,226	7,226

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

This exercise also shows that in the regressions, with income variable as an extra control, there is a similar significant and positive correlation between the instrument variable and migration in the first stage and large F statistics on the excluded instrument.

However, one may ask why would people migrate? If the effects of migration on time use of the left behind are not significant, then what is the benefit of migration?

First, the findings in this paper do not suggest that migration has any effect on the left behind. Migration impacts the time use and labour supply of the left behind through two channels, the income effect and the substitution effect. These two effects of migration work together and in different directions; thus, they make the net effect of migration on the time use of the left behind non-significant and, not statistically different from their

non-migrant counterparts.

Second, the no effect results on time use and labour supply do not mean that migration does not make the family members better off. The welfare of the left behind can be measured in various aspects, whether socially, economically, psychologically or physically. Other aspects of the economic welfare of the left behind may also be important; the left behind members could have more consumption and nutrition, a greater ability to invest in the school, to make purchases of durable goods and agricultural machines, etc. These aspects of welfare closely relate to the quality of life and satisfaction of the left behind children and women. They could be very interesting empirical research topics in the future.

Table 2.9: FEIV with income variable: the impact of husband's migration on spouse's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.038**	0.038**	0.038**	0.038**	0.038**	0.038**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
First stage F-statistics	4.33	4.33	4.33	4.33	4.33	4.33
Prob > F	0.038**	0.038**	0.038**	0.038**	0.038**	0.038**
Second Stage						
Husband migration status	0.824	2.025	0.322	2.273	-0.148	2.870
	(0.744)	(1.602)	(0.453)	(1.554)	(0.741)	(5.759)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	6.64**	6.64**	6.64**	6.64**	6.64**	6.64**
Weak identification test	4.33**	4.33**	4.33**	4.33**	4.33**	4.33**
Weak instrument robust inference	2.57	3.97**	0.98	6.57**	0.06	0.42
Durbin-Wu-Hausman test	4.33**	4.12**	4.63**	3.25*	5.06**	5.68**
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	1,792	1,792	1,792	1,792	1,792	1,792
R square	0.622	0.495	0.388	0.075	0.624	0.605
N	5,266	5,266	5,266	5,266	5,266	5,266

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

2.6 Robustness

To make the main findings of the paper more convincing, three robustness checks are conducted; (1) estimating the main equation by sub-sample, according to the gender of the children, (2) finding the impact of migration on the time use of total activities and self-employment of left behind children and spouse, (3) estimating the main equation with robust standard errors cluster at village level.

2.6.1 Subgroup Results

Left behind children could behave differently to their father's migration depending on their gender. The estimation results from different gender groups of the children are presented in Tables 2.10 and 2.11, where Table 2.10 shows the boys' results and Table 2.11 presents the girls' results.

The first stage regressions indicate a strong and positive correlation between the instrument and migration. The F statistics on the excluded instrument are relatively large (7 for boys' outcomes and 5.4 for girls' results). However, as the F statistics are under 10, the results of this robustness check need to be interpreted with attention.

When estimating the impact of a father's migration on boys and girls separately, the results are non-significant for all the activities, except the boys' time use in wage labour. This makes the main results more persuasive, and more convincing to make the statement that the time use and labour supply of the left behind children are not statistically different from their counterparts, the income and substitution effects of migration work together and cancel each other out for subgroups of the sample.

The main results are robust to the division of the sample. By comparing the results in Tables 2.10 and 2.11, it is clear that the girls' time use in agriculture and housework is more sensitive to the father's migration; however, the impact of migration on children's wage labour is driven by boys. The boys would increase participation in wage labour by around 70% and increase the hours per day in wage labour by as much as 5 hours,

while the net impact of the fathers' migration on girls' wage labour is not significant.

One may argue that this finding is contradictory to the common value of preferring a son in rural China. Firstly, it's true that individuals in rural China favour sons. Sons are traditionally regarded as heirs and labour reserves for the family, and thus rural family may allocate more resources to boys (for instance, education investment). However, in this analysis, the special situation is that when the father migrates, the children and mother are left behind. It's expected that when the father leaves family, the mother has stronger bargaining power and the allocation of family resource would shift to girls who are always treated unfairly (Antman (2011a)). This is also in line with the empirical findings that father's migration favours the welfare of their daughters.

The results support the main finding in the sense that income effect and substitution effect in most activities are trade off for boys and girls. Another finding states that only boys would take up the wage labour job if the family did not receive remittances immediately after the father's migration and faced financial difficulties during that short period. This finding is consistent with the subgroup results of Antman (2011a).

From the subgroup study of children's outcome, the results indicate supportive evidence of the main findings. They also show that the impact of the father's migration on child labour is mainly driven by boys.

Table 2.10: FEIV estimation: the impact of father's migration on boy's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.060***	0.060***	0.060***	0.060***	0.060***	0.060***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
First stage F-statistics	6.97	6.97	6.97	6.97	6.97	6.97
Prob > F	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Second Stage						
Father migration status	0.313	1.762	0.029	-0.209	0.685**	5.150**
	(0.354)	(1.748)	(0.538)	(0.611)	(0.347)	(2.610)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	10.98***	10.98***	10.98***	10.98***	10.98***	10.98***
Weak identification test	6.97***	6.97***	6.97***	6.97***	6.97***	6.97***
Weak instrument robust inference	1.76	2.65	0.00	0.23	17.10***	18.73***
Durbin-Wu-Hausman test	4.16**	4.98**	2.10	2.36	9.09***	8.71***
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	1,442	1,442	1,442	1,442	1,442	1,442
R square	0.424	0.414	0.459	0.399	-0.488	-0.499
N	3,926	3,926	3,926	3,926	3,926	3,926

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

2.6.2 Self-employment and Total Time Use

The time use on agriculture, housework and wage labour of the left behind children and spouse is not influenced by the migration of the male household head. Moreover, the time use of the left behind on the family business can also be considered. This business can be a store for handicrafts, groceries, etc. The migration of the male household head would generate more non-labour income for the left behind family to start or enlarge the family business, or it may be the case that the left behind family member would need to compensate for the lost labour due to migration. The income effect and substitution effect of migration then work in the same direction on the time use in the family busi-

ness. However, the income effect of migration can also lead to a decrease in the time use in the family business as the higher reservation wage of the left behind, and they can maintain the utility level with less labour supply.

Table 2.11: FEIV estimation: the impact of father's migration on girl's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.060**	0.060**	0.060**	0.060**	0.060**	0.060**
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
First stage F-statistics	5.37	5.37	5.37	5.37	5.37	5.37
Prob > F	0.021**	0.021**	0.021**	0.021**	0.021**	0.021**
Second Stage						
Father migration status	0.841*	2.908*	-0.952	-1.909	0.163	1.040
	(0.502)	(1.743)	(0.769)	(1.450)	(0.204)	(1.747)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	8.77***	8.77***	8.77***	8.77***	8.77***	8.77***
Weak identification test	5.37**	5.37**	5.37**	5.37**	5.37**	5.37**
Weak instrument robust inference	11.14***	15.50***	3.88**	6.42**	1.51	0.87
Durbin-Wu-Hausman test	5.43**	5.97**	3.16*	4.45**	2.83*	2.40
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	1,230	1,230	1,230	1,230	1,230	1,230
R square	-0.011	0.168	0.338	0.187	0.405	0.416
N	3,300	3,300	3,300	3,300	3,300	3,300

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

This section presents the impact of migration on the time use of the left behind in the family business and also the total time use in all sorts of activities, including agriculture, housework, wage labour and family business. Table 2.12 shows the fixed effects and instrumental variable estimation for the children's outcomes. First, the first stage regressions of the father's current migration status are positive and significant (0.058), which shows the correlation of the IV with migration. The first stage F-statistic are 12, which is a signal that the IV is strong. The second stage shows the migration of the father has no significant impact on the self-employment or total work of the

left behind children, neither in participation nor hours spent. This result is a strong supportive evidence to the main finding of this paper that the net impact of the father's migration on the time use of left behind children is not significant.

In addition to the outcomes of the left behind children, the same estimation results for the family business and total work of the left behind spouse are presented in Table 2.13. The first stage F statistics is significant, but is under 10; however, the first stage coefficient of the instrumental variable is 0.038, which is positive and significant. The interpretation of results of the left behind spouse, hence, needs some attention. The focus of the results is that the coefficients on the migration of husband are not significant for the self-employment and total work. This finding also supports the robustness of the main finding; the total effect of migration on the left behind spouse is not significant.

The results of the time use of business and total work indicate that for the left behind children and spouse, the income effect and substitution effect of the migration of the male household head are traded off, which supports the main finding.

2.6.3 Cluster at Village Level

This robust check considers the fact that rural residents, especially left behind children and women, impact others in the rural villages. The time use behaviour of the left behind children and women may influence, or get influenced by, the behaviour of those in non-migrant families of the same village. Thus, this robustness check uses the cluster robust standard errors clustered at the village level, to allow for any arbitrary correlation between the rural residents of the same village. The regression results with cluster robust standard errors for left behind children and women are presented in Tables 2.14 and 2.15 respectively. Although the first stage F statistics are relatively large (7 for children outcome and 5 for women outcome), they are not larger than 10 to show the power of the IV. Hence, the results of the IV regressions would need to be interpreted with caution. However, the first stage coefficients on the instrumental variables are positively large and significant (0.06 and 0.04 for children and women outcomes respectively). The

fixed effects and instrumental variable with clustered standard errors allow for arbitrary correlation between individuals within the village. This is a strong assumption that may lead to a decrease in F statistics.

Table 2.12: Robustness check: the impact of father's migration on children's time use and labor supply

	family business		total work	
	participation	hours	participation	hours
First Stage				
Wage Difference	0.058***	0.058***	0.058***	0.058***
	(0.017)	(0.017)	(0.017)	(0.017)
First stage F-statistics	12.00	12.00	12.00	12.00
Prob > F	0.00***	0.00***	0.00***	0.00***
Second Stage				
Father migration status	-0.180	-0.813	-0.208	3.544
	(0.101)	(0.587)	(0.440)	(2.227)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	19.12***	19.12***	19.12***	19.12***
Weak identification test	12.00***	12.00***	12.00***	12.00***
Weak instrument robust inference	13.46***	8.80***	0.37	5.67**
Durbin-Wu-Hausman test	4.12**	3.74*	2.18	3.17*
Individual Controls	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes
Number of fixed effects	2,667	2,667	2,667	2,667
R square	0.194	0.272	0.500	0.378
N	7,226	7,226	7,226	7,226

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV estimation.

The coefficients for the father/husband migration status for all the activities are not statistically significant, also for the children's wage labour which shows some significance

in the main result. This robustness study does support the main finding that migration has no significant impact on the time use of the left behind children and women. It also shows that the impact of father migration on the child wage labour vanishes when the correlation within the villages is allowed.

Table 2.13: Robustness check: the impact of husband's migration on spouse's time use and labor supply

	self-employment		total work	
	participation	hours	participation	hours
First Stage				
Wage Difference	0.038**	0.038**	0.038**	0.038**
	(0.018)	(0.018)	(0.018)	(0.018)
First stage F-statistics	4.33	4.33	4.33	4.33
Prob > F	0.038**	0.038**	0.038**	0.038**
Second Stage				
Husband migration status	-0.612	-2.442	0.541	3.647
	(0.600)	(1.724)	(0.384)	(6.346)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	6.64**	6.64**	6.64**	6.64**
Weak identification test	4.33**	4.33**	4.33**	4.33**
Weak instrument robust inference	2.26	5.05**	7.69***	5.56**
Durbin-Wu-Hausman test	2.55	3.44*	3.53*	3.15*
Individual Controls	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes
Number of fixed effects	1,792	1,792	1,792	1,792
R square	0.478	0.718	0.037	0.483
N	5,266	5,266	5,266	5,266

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust standard errors in parentheses. FEIV.

2.7 Conclusion

There are many left behind children and women in rural China, who are separated from the male household head, yet this issue has been studied less in the economics literature.

Table 2.14: FEIV estimation with cluster: the impact of father's migration on children's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.060***	0.060***	0.060***	0.060***	0.060***	0.060***
	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)	(0.022)
First stage F-statistics	7.32	7.32	7.32	7.32	7.32	7.32
Prob > F	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
Second Stage						
Father migration status	0.532*	2.329*	-0.527	-1.189	0.428*	3.095*
	(0.312)	(1.508)	(0.511)	(0.901)	(0.230)	(1.755)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	6.52**	6.52**	6.52**	6.52**	6.52**	6.52**
Weak identification test	7.32***	7.32***	7.32***	7.32***	7.32***	7.32***
Weak instrument robust inference	7.74***	7.91***	2.15	5.75**	11.64***	8.56***
Durbin-Wu-Hausman test	4.05**	3.36*	4.76**	2.79*	6.21**	5.11**
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of cluster: villages	147	147	147	147	147	147
Number of fixed effects	2,667	2,667	2,667	2,667	2,667	2,667
R square	0.288	0.314	0.418	0.280	0.054	0.098
N	7,226	7,226	7,226	7,226	7,226	7,226

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. 2662 individual fixed effects nested within cluster
Cluster robust standard errors in parentheses. FEIV estimation.

The empirical evidence on the impact of migration on the left behind children and women in rural China, especially their time use and labour supply behaviour is limited. Migration has two effects on the left behind family members, namely the income effect and substitution effect. Studies in international or internal migration with different countries show diverse results; while some find that there is a larger income effect, some

identify a greater substitution effect. This paper contributes to the migration literature by answering the open empirical question - what is the total effect of migration on the time use of the left behind children and spouse? This paper indicates that the income and substitution effects cancel each other to make the total effect of migration on the left behind statistically insignificant.

Table 2.15: FEIV estimation with cluster: the impact of husband's migration on women's time use and labor supply

	agriculture		housework		wage labour	
	participation	hours	participation	hours	participation	hours
First Stage						
Wage Difference	0.041**	0.041**	0.041**	0.041**	0.041**	0.041**
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
First stage F-statistics	4.95	4.95	4.95	4.95	4.95	4.95
Prob > F	0.028**	0.028**	0.028**	0.028**	0.028**	0.028**
Second Stage						
Husband migration status	0.748	1.876	0.285	2.129	-0.201	2.125
	(0.781)	(1.656)	(0.439)	(1.464)	(0.715)	(5.888)
Overidentification test	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified	exactly identified
Underidentification test	4.76**	4.76**	4.76**	4.76**	4.76**	4.76**
Weak identification test	4.95**	4.95**	4.95**	4.95**	4.95**	4.95**
Weak instrument robust inference	1.71	3.36*	0.73	3.76*	0.10	0.19
Durbin-Wu-Hausman test	3.96**	3.28*	4.46**	3.06*	4.13**	4.76**
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of cluster: villages	147	147	147	147	147	147
Number of fixed effects	1,792	1,792	1,792	1,792	1,792	1,792
R square	0.641	0.523	0.402	0.013	0.619	0.614
N	5,266	5,266	5,266	5,266	5,266	5,266

CHNS 1997-2011, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. 1787 individual fixed effects nested within cluster
Cluster robust standard errors in parentheses. FEIV estimation.

This paper not only contributes to the limited literature in this area, but also to the methodology of the existing studies. The past literature either used the pooled cross-sectional aspect of the panel date, or the simple linear regressions or the instrumental variable approach, but the IV can be strongly correlated with the outcomes of the left

behind. Understanding the determinants of rural to urban migration in China helps to construct IV in the empirical model. The FEIV strategy deals with the time variant and invariant sources of endogeneity of migration.

In this study, the participation and time use on three types of activity, i.e. wage labour, housework and agriculture, are studied for both the spouse and children left behind. The total time use and the time use in self-employment are also examined for the robustness check. The study identifies that the total effect of migration on the time use and labour supply of the left behind children and spouse is non-significant, suggesting that the income and substitution effects of migration cancel out.

One exception from the main findings is that father's migration increases the boy's time use in wage labour. This may be due to the delay in the remittances from migrants that would lead to a delay in the income effect of migration. However, in the case of within village correlation, this impact vanishes and makes the results similar to the main findings. The increase of the boy's time use in wage labour is also expected to minimise when the left behind receives the remittances a little after migration. Thus, the income effect would trade off the substitution for the boy's time use in wage labour. Considering the immediate financial hardship that may occur after migration, it may impact the schooling and health of the left behind boys.

The income effect results from remittances generated from migration, which lead to a decrease in the time use and labour supply of the left behind. However, the substitution effect works in the opposite direction; the left behind may have to increase the time use to compensate for the lost labour within the household due to migration. Thus, the two effects cancel out and this leads to the non-significant coefficients, which have been identified in this paper.

Even though the time use of the left behind is related to many critical welfare aspects like health and education, the findings of this paper do not indicate that migration makes the left behind better off or worse. The fact is that other aspects of the welfare of the left behind children and spouses, such as the consumption, nutrition and health, education and cognitive skills, etc., may be affected more than of the children and spouses in

non-migrant households. It is recommended that a further study can be conducted to examine and evaluate this aspect of migration effect.

3

Return Migration and Consumption Behaviour

3.1 Introduction

Return migration, both internally and internationally, has been attracting excessive academic focus and is a topic of public debates over the past few years. Moreover, research on return migration has been steadily increasing (Clemens et al. (2014)). However, there is a lack of evidence regarding the impact of return migration beyond the labour market, especially on return migration in the context of internal rural to urban migration, in developing countries.

Amidst the growing literature on the economics of return migration, the maximum focus has been given to the theoretical or empirical analysis of determinants of return and labour market outcomes of return migrants. Return migrants accumulate savings and human capital and their return to the place of origin has multiple impacts on the economics and labour markets, in both the destination and the place of origin (Wahba

(2014)).

This paper intends to answer the following question: How does migration impact the consumption patterns, especially the consumption of smoking and drinking when the migrants return? To achieve this, a panel survey from China is used for examining novel empirical questions like, are return migrants more likely to drink and smoke than non-migrants and do return migrants consume more tobacco and alcohol than non-migrants. By comparing return migrants and non-migrants, this study addresses the selection problem, which may bias estimation of the effect of return migration on drinking and smoking behaviours. The before and after migration behaviour is also examined in order to study how migration leads to a change in the smoking and drinking pattern.

This study focuses on China, which has seen massive rural to urban migration and provides an excellent source of research in the economics of migration. As stated in Meng (2012): "From the late 1990s up to the present, the number of rural migrants increased by 100 million to 145 million ", and in Zhao (2002): "Return migration is an integral part of the rural to urban labour migration in China". More interestingly, smoking and drinking behaviour is common in China. According to the World Health Organisation (WHO), each year approximately 1 million deaths in China are caused due to excessive tobacco consumption, while the death rate attributed to alcohol consumption in China is around 65% among the health-related deaths.

Theoretically, return migrants are expected to acquire more savings and human capital and alter their behaviour towards employment, investment and consumption. Another channel of behaviour change is that return migrants, who are influenced by the social norms in urban areas and adapt to the lifestyle of the migration destination, can potentially transmit this acquired behaviour and norms to the place of origin. The tobacco and alcohol consumption are a critical aspect as they may have an enormous effect on individuals, households and communities. As smoking and drinking are hazardous and addictive, returnees' behaviour can affect their own health and employment and negatively influence the family members, especially children and elderly parents. This can lead the household to a risky financial or health situation.

The contribution of this paper is to measure the impact of return migration on the smoking and drinking of returnees by addressing the selection problem, as well as by comparing the before and after migration behaviour change by using the national panel survey data. To best of the knowledge, no previous study has been carried out till now, regarding the empirical investigation of the tobacco and alcohol consumption behaviour of the return migrants. This work contributes to the rapidly developing research on return migration, and particularly focuses on drinking and smoking behaviour using the panel data, which comprises the information on migration, smoking and drinking behaviour, as well as on other community, household and individual characteristics. Hence, this paper contributes to the emerging research on the transfer of social norms through return migration.

Self-selection bias is an important issue of migration studies, not just for the selection of migration, but also the selection of return from the migrants (Wahba (2015)). However, in the case of China, migration is temporary because of the "Hukou" system. As stated in Démurger and Xu (2011): "These administrative barriers to permanent settlement in cities tend to make rural migrants return to their home community." This paper studies the smoking and drinking behaviour of return migrants; thus, there is no reverse causality effect problem, as smoking and drinking behaviour is unlikely to have a direct effect on return migration decisions (Some argue that migrants return because they prefer consumption in their home country. In the case of international migration, this can be a good explanation for the return. However, in the case of internal rural to urban migration, the marginal utility of consumption are the same). Nonetheless, migrants are selected, and ignoring the selection problem in this empirical analysis will lead to bias in the results.

This paper uses the panel data China Health and Nutrition Survey (CHNS) from the year 1997 to 2011. Considering the selection problem and controlling for social and economic characteristics at different levels, this study finds that male return migrants smoke and drink more than non-migrants. Return migration has a strong effect on male outcomes. This paper corrects the selection bias using a unique instrumental variable. Furthermore, Difference-in-Difference (DiD) strategy is applied for comparing the be-

fore and after migration behaviour changes in smoking and drinking and the results are consistent with the instrumental variable strategy.

The rest of the paper is organised as follows: Part 2 reviews the literature and introduces the background on migration and return in China. Part 3 presents the data and statistics, followed by empirical models and results of instrumental strategy in Part 4. Part 5 presents the construction of DiD and the before and after migration changes in smoking and drinking. Part 6 conducts robustness checks, and Part 7 concludes the paper.

3.2 Literature

Academic effort to explain and model return migration lasts decades (see Wahba (2014) and Dustmann and Görlach (2016) for a review). Due to the lack of data, empirical evidence on return migration and its impact has been insufficient. According to Wahba (2014), the impact of return migration on sending areas has mainly been examined via the following channels: investment behaviour and entrepreneurship of returnees; human capital and wage of returnees; transfer of social norms and political ideas.

Despite the existing and growing empirical or theoretical work on the investment, saving and employment behaviour of return migrants, the literature on the consumption behaviour of returnees is exiguous. Migrants would accumulate assets in migration destination, and then finance higher consumption upon return (see Dustmann and Görlach (2016)). However, no empirical work has studied the consumption behaviour of the return migrants.

Research relevant to this study includes the emerging literature investigating the impact of return migration on the transmission of social norms. Bertoli and Marchetta (2015) shows that Egyptian return migrants adjust their fertility choices to the norms in migration destination countries. They find that return migrants who have past migration experience in another Arab country, have a significantly larger number of children when

compared to the stayers. Similarly, in this paper, the migrants returning to their home villages may change their smoking and drinking behaviour according to the norms that prevailed in the cities, where the migrants worked.

Compared to the literature on international migration in other developing countries, research on the impact of return migration in rural to urban migration in China is limited. Some existing work focuses on the determinants of return. Démurger and Xu (2013) concludes that left-behind children pull their parents back to the village. Similarly, Zhao (2002) states that return migrants tend to be older, married, better educated and with a left behind spouse. Other studies focus on the employment and investment behaviour of the return migrants. Démurger and Xu (2011) finds that return migrants are more likely to be self-employed and this probability increases with the savings and frequency of job changes of returnees. Ma (2002) finds that human capital accretion during migration reinforces the mobility of social capital, and enhances the returnee's entrepreneurship; Ma (2001) finds that human capital acquired from migration strongly facilitates the occupational change of return migrants.

Some papers study the consumption behaviour in the context of migration. Gibson et al. (2011) uses consumption data on rice, fruits, fats, milk and meat. They find that children who migrate to New Zealand enjoy a better diet and are in better health, while the children left behind in Tonga when other family members migrate, shift to lower-cost consumption and worse health.

One exception that mentions the consumption behaviour of return migrants in China is Zhao (2002). By comparing the ownership of consumer durable goods, housing and production assets of migrants, non-migrants and returnees, it is found that returnees contribute more in all types of assets as compared to non-migrants and migrants, and invest more heavily in a production machine. The drawback of this estimation is that she doesn't consider any types of selection problem in the empirical work.

Since the "Open and Reform" policy in the early 1980s, migration and return between rural and urban areas in China have been a vital social and economic phenomenon. The reasons for returning to villages can be complex. The most important one could

be the "Hukou" system, which restricts migrants and their families to get access to employment, housing, pensions, education, health and other public services that only urban residents can avail (see Meng and Zhang (2001) for a formal discussion on the Chinese "Hukou" System).

Over the years, the authority has gradually relaxed the rural to urban mobility. Some migrants successfully obtained urban residential status. Nonetheless, the obstacles remain that the rural migrants and their families are still in disadvantaged position, as compared to the urban residents in various aspects. Meng and Zhang (2001) finds that rural migrants obtain lower-end jobs and earn less, and claim that "less than 4% of rural migrants have a pension provided by their employers or the governments, compared with 90% for urban residents. The percentages of migrants and urban residents who enjoy medical insurance provided by employers or government are 9% and 77%, respectively". For decades, rural to urban migration has been massive and steadily growing. Due to the household registration system in China, migrants from rural areas working in cities are regarded as temporary rural workers by the authority.

Besides the contribution to the literature of return migration, the consumption behaviour of return migrants is an essential research question in multiple aspects. For policy makers, it's important to understand the consumption behaviour of return migrants. The Chinese central government wants to boost economic growth by stimulating domestic consumption, instead of the investment and export it has concentrated on for the past 3 decades (Song et al. (2011)). As a traditional saving country ("the saving rate in China exceeds that of nearly every other country" Curtis et al. (2011)), China may embrace economic growth if households would consume and invest more, instead of saving too much. Compared to non-migrant farmers, migrants have accumulated more savings and skills. Once they return, they may, on one hand, have more incentive and stronger power to consume; and on the other hand, they may have more precautionary saving. It's vital and interesting to understand their behaviour in consumption. At the individual and household levels, return migrants act as a bridge between rural and urban areas in China. They may adopt drinking and smoking behaviour and bring it back to the villages. In addition, with more disposable incomes they have an incentive to consume

more in the aspect of smoking and drinking. Psychologically, they may want to spend more, or purchase branded goods to show off to their old friends and neighbours in the village. On the one hand, smoking and drinking are an unhealthy behaviour for individuals, family and villages. On the other hand, the policy makers would encourage return migrants to spend rationally, for instance, to start a micro-enterprise, invest in education and training, allocate resources for productive machines and durable goods, instead of drinking and smoking, or on other behaviours, such as gambling, buying brand goods, etc.

3.2.1 Drinking and Smoking in China

Smoking and drinking are popular behaviours in China in both urban and rural areas, especially among males. According to the data from the World Health Organization (WHO)[1], there are more than 300 million smokers in China, which is near $1/3$ of the world's total population. Also, one in every three cigarettes smoked in the world is consumed in China and the smoking rate of males and females in 2010 were 52.9% and 2.4% respectively.

"For Chinese residents aged 15-69 years in 2007, about 55.6% males and 15.0% females drink" by Li et al. (2011). "Along with rapid economic growth in China, there has been a striking increase in alcohol consumption, greater than in most other parts of the world", as described in Jiang et al. (2015).

Table 3.1 shows the tobacco and alcohol consumption behaviour of males and females in rural areas for the year 1997 to 2011. The smoking and drinking rates for rural residents from CHNS data coincide with the WHO data. Male smoking rates range from 60% to 70%, while smoking rates for female decreased over the years from over 6% in 1997 to below 5% in 2011. The rates from this data are larger than the WHO data (52.9%, 2.4%) because sample in this paper is restricted to the 15 to 70 age group [2], while the WHO data are smokers among the entire population. Around 50% to 60%

[1] WHO report: http://www.wpro.who.int/china/mediacentre/factsheets/tobacco/en/

[2] The compulsory education in China (primary school and junior high school education) ends at age

rural males are alcohol consumers, while for females the proportion is about 5% to 8%. The rural males' drinking rate is the same as the WHO data.

Table 3.1: Smoking and drinking by gender and year

	Male				Female			
	Smoker		Drinker		Smoker		Drinker	
Year	Freq	Rate	Freq	Rate	Freq	Rate	Freq	Rate
1997	2,178	72.29%	1,797	59.64%	182	6.45%	215	7.62%
2000	2,510	72.09%	1,895	54.42%	213	6.19%	227	6.60%
2004	2,148	66.65%	1,704	52.87%	176	5.33%	182	5.51%
2006	2,009	64.49%	1,684	54.06%	172	5.20%	198	5.98%
2009	2,024	64.75%	1,754	56.11%	157	4.85%	205	6.34%
2011	1,756	64.32%	1,525	55.86%	139	4.70%	189	6.39%

CHNS 1997 to 2011, 15 to 70 years old rural residents.

Tobacco and alcohol are consumed at major family events, social interactions and also for daily consumption. Drinking and smoking can be very harmful to health and family. In China, "one person dies around every 30 seconds because of tobacco use" (WHO). Lancet (2014) states that "Annual health costs of an estimated ¥ 41 billion have been attributed to the major smoking-related diseases." Besides that, exposure to second-hand smoke can be harmful to family members, such as children, elderly parents and pregnant women. Drinking alcohol, especially heavy drinking, can lead to serious issues. WHO Bulletin reported that [3]:"As in many other countries, in China, excessive drinking has shown an association not just with health-related harm, but also with social harm, specifically traffic accidents, crime, child abuse, domestic violence and injuries of all types."

Given the popularity and adverse effect of tobacco and alcohol, it's important for

15, thus the population aged over 15 years are included in this study. Given the span of the data is about 25 years, and high migration and return rates at middle age, it's reasonable to examine the population up to 70 years old sample. For instance, someone migrates at 45 in 1997 and would be 70 in the year 2011.

[3]WHO bulletin: http://www.who.int/bulletin/volumes/91/4/12-107318/en/.

academia and policy makers to better understand this consumption behaviour for return migrants, who have relaxed budget constraints and links the rural and urban China.

3.3 Data

The data this paper uses is a panel from the China Health and Nutrition Survey (CHNS) conducted by the Carolina Population Center at the University of North Carolina and Chinese Center for Disease Control and Prevention. This data includes a household survey, individual survey, nutrition & physical examination and community survey for the year 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011. The data has about 4400 households with $26,000$ individuals across 9 provinces in China.

Variables regarding smoking and drinking in this study are derived from the questions on this behaviour of individuals, such as: Do you drink beer/liquor/grape wine? How much do you drink each week? Have you smoked cigarettes? Do you still smoke cigarettes now? How many cigarettes do you smoke per day? The incidence and intensity of smoking and drinking behaviour are studied in this paper though it focuses on the behaviour of the return migrants and does not show how or when smoking and drinking is harmful. Controlling for different income variables in the model would show that the behaviour change in smoking and drinking is not simply because of an income effect.

Migration variable is derived from the question of whether the household member is currently living in the household. If no, is he seeking employment elsewhere? Questions about migration are asked in the survey since 1997, hence the 1997 and the following waves are used. Return migration is constructed by previous migration status. For instance, if one migrates in 2004, and does not migrate in the next wave 2006, then he or she is identified as a return migrant in 2006. By this identification, the migration history of individuals can be tracked, and rural residents are divided into three groups: returnee, migrants and non-migrants. However, it's impossible to capture the data on all the return migrants because of the temporary property of migration in China. There are both migration and return between the waves. This paper has a lower measurement

of the migration and return migration due to unclear migration history records in the survey.

This minor drawback shouldn't lead to serious bias in the analysis. This paper relies on variation in the change of return migration status and compares the before migration and after migration behaviour change for migrants and non-migrants. Information on migration, return or stay between panels is missing randomly and asymmetrically missing in this data, when the survey implemented random sampling method. The measures of return migration correspond with the published work using other Chinese data. Démurger and Xu (2011) have 9% return migration from a household survey in Wuwei in 2008. Zhao (2002) uses a survey from the Chinese Ministry of Agriculture to find that the return rate is about 8% in 1999.

Figure 3.1 shows the number of migrants and the rates of migration over the survey years in CHNS. Male migration rates increase rapidly from around 8% in 1997 to 25% in 2006 and then remain relatively steady. Female migration rates are lower, but also increase quickly from about 6% in 1997 to 15% in 2006 and then stay relatively stable. The number of male migrants increases from about 300 in 1997 to more than 1000 after the year 2006. For females, the number is around 200 in 1997 and over 600 after the year 2006.

Figure 3.1: Migrants and migration rates

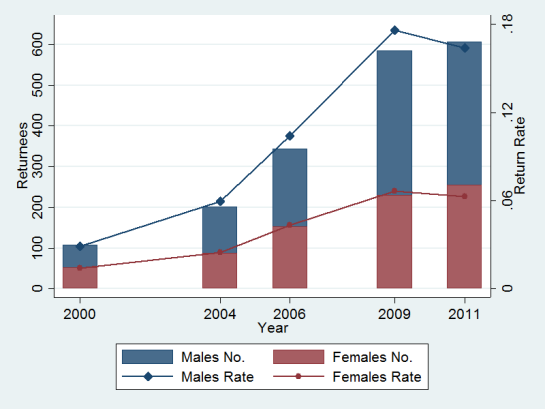


Figure 3.2: Returnees and return rates

CHNS 15 to 70 years old rural residents. Whole bar (blue + red) for male, red for female.

Figure 3.2 is about the number of return migrants and return rates. Returnees are non-migrants with previous migration experience, and return rates are the percentage of return migrants in the population of 15 to 70 age group. There are no observations of return migrants in the first panel 1997. For males, return rates increase steadily from 3% in 2000 to around 18% in 2009. For females, the rates are about 1.5% in 2000 and over 6% in 2009. The number of male returnees is around 100 in 2000 and goes above 600 in 2011. The number of female returnees in 2000 is around 50 and it is over 250 in 2011.

Even though not specially designed for migration study, the data contains abundant information on migration and return migration as displayed in the Figures 3.1 and 3.2. For more information on migration and return in the data, Tables 6.4 and 6.5 in Appendix 6.3 are presented to describe the migration times of individuals and statistics on those return migrants, who migrate again in the following year.

Apart from the migration patterns, the timing of the survey is also a challenge. One may argue that Chinese people consume tobacco and alcohol much more during the Chinese Spring Festival (when migrants always return to their villages for celebrating the traditional festival). Table 6.6 in Appendix 6.3 shows that time points of the survey are mostly distributed in August to December in the year, not January to March for the Spring Festival. This suggests that the concern about the timing of the survey in this study is not a problem for this analysis.

The paper cannot capture the fact that all migrants have to return to their villages. The highly temporary and repeated property of rural to urban migration in China is not fully captured through statistics from this data. After all, it is discovered that return migration rates increase rapidly for both males and females, and the return migration rates increase with the migration rates. Moreover, from Table 6.5 in Appendix 6.3, the migration rates of returnees are relatively high, with around 30% returnees migrating in the following survey year. This indicates the highly temporary and circular character of rural to urban migration in China. This is also an important variation that this paper would exploit.

Besides migration & return and smoking & drinking consumption, the data provides information on the economic and social backgrounds of individuals and households, which are presented by the return status in Table 3.2. A set of village indexes, newly created by the survey, is also used to control for the village characteristics.

In 37,754 sample observations, there are 2,286 observations with return migration status of 1 and 35,468 of 0. The sample includes 9,869 individuals, 4,934 males and 4,935 females. 1517 individuals experienced a return migration status change across the years, which is about 15% of the whole sample size. There are more observations of male returnees than females returnees, 1,655 and 631 respectively. 4,931 observations are discarded as observed only once. On average, returnees are about 4 years younger than non-migrants and they are more educated especially female returnees. Returnees are more likely to be single rather than married, as compared to the non-migrants. The household and individual income of return migrants are significantly higher than that of non-migrants. This is plausible as migrants accumulate savings and return with more assets and human capital in order to earn a higher income.

The physique of migrants are better; they are significantly taller and heavier than non-migrants. This paper builds an identification strategy based on this fact. Two health related indexes are included to control for the issue that one may argue healthy people tend to drink more and smoke more. "Sick" is an indicator of "whether the person gets any types of illness in the past 4 weeks at the time of the survey". There is no difference between the two groups for the probability being sick. "Med Insurance" is an indicator of "does this person have any kind of medical insurance". Significantly, both male and female returnees have a higher probability of having medical insurance.

The average smoking incidence for returnees is 16% higher than that of non-migrants, and the drinking incidence is 14% higher for return migrants. It's also interesting to note the drinking and smoking amounts for returnees are both higher. Male returnees have significantly larger smoking and drinking incidences. Though the amount of smoking is more, the amount of drinking is lesser. Female returnees and non-migrants, on the other hand, have no difference in the average smoke amount and drink incidence. In the fol-

	Whole Sample			Males			Females		
	Non-migrant	Returnee	Diff	Non-migrant	Returnee	Diff	Non-migrant	Returnee	Diff
Age	42.91	38.84	4.10***	42.30	39.51	-2.79***	43.54	37.09	6.45***
Education	0.71	0.83	-0.12***	0.85	0.86	-0.01	0.59	0.75	-0.16***
Marriage	0.84	0.80	-0.04***	0.80	0.81	-0.01	0.88	0.79	0.09***
HH Income	21472	30377	-8905***	21365	31123	-9758***	21570	28412	-6842***
Ind Income	6302	9706	-3403***	7512	10798	-3286***	5184	6839	-1655***
Height	161	164	-3***	166.4	166.8	-0.4***	156	157	-1***
Weight	59	61	-2***	63.0	63.9	-0.9***	56	55	-1***
Sick	0.13	0.13	-0.00	0.11	0.13	-0.02**	0.14	0.12	-0.02
Med Insurance	0.43	0.68	-0.25***	0.43	0.69	-0.26***	0.43	0.67	-0.24***
Smoke	0.35	0.51	-0.16***	0.67	0.69	-0.02*	0.06	0.03	0.03***
Smoke Amount	4.94	7.52	-2.58***	9.88	10.27	-0.39**	0.38	0.29	0.09
Drink	0.30	0.44	-0.14***	0.55	0.60	-0.05***	0.06	0.05	0.01
Drink Amount	1.74	2.31	-0.57***	3.21	3.16	0.25***	0.20	0.08	0.12***
N	35,468	2,286		17,034	1,655		1,8434	631	

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. CHNS 1997 to 2011, 15 to 70 years old population.

Table 3.2: Descriptive statistics by return status

lowing sections of the paper, the study will examine whether or not return migrants are more likely to smoke and drink, and whether they consume more tobacco and alcohol than non-migrants, addressing the self-selection of migration and comparing the before and after migration behaviour change.

3.4 Empirical Model

To examine the effect of return migration on tobacco and alcohol consumption patterns, consider the linear regression model of the following form:

$$Y_{i,t} = \alpha + \beta R_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \qquad (3.1)$$

with $\varepsilon_{i,t} = c_i + u_{i,t}$.

$Y_{i,t}$ is the outcomes of interest: the smoking and drinking of rural individual i at time t. $Y_{i,t}$ can be a binary variable indicating for the individual smoke/drink or not and can also be the amount of smoking and drinking. $R_{i,t}$ is an indicator of the return migration status of individual i at time t. It is 1 if individual i is a return migrant at time t, otherwise is 0. $X_{i,t}$ is a set of control variables at the individual, household and village levels. For details about the explanation and summary statistics of these controls, check Tables 6.8 and 6.9 in Appendix 6.6. c_i is the individual fixed effects that can capture individual time-invariant unobservable factors that may lead to the correlation of $R_{i,t}$ with $u_{i,t}$. α is the constant term, β and γ are the parameters.

If the self-selection problem is not considered at this moment, pooled OLS and panel fixed effects estimation results of Equation 3.1 are presented in Table 3.3. Controls at individual, household, village levels and time dummies are included. For individual characteristics, the age, square of age, marriage status, education and health related index are modelled. At the household level, the controls are measures of household assets, such as ownership of phone and type of energy in use. A set of village characteristics is controlled for (See Table 2.2 for details), such as the village education level, urbanisation index, population density and economic score. The standard errors are robust to

heteroskedasticity and are clustered at the village level.

From Table 3.3, the coefficients on return migration (β) in the pooled OLS models are significant and positive. As always, the OLS regressions are not to be trusted. In the linear fixed effects models, the coefficients on the smoking behaviour are significant, but for the drinking outcomes they are not. The fixed effects regression model can deal with the time-invariant individual heterogeneities that may lead to a bias of the estimation. However, it is not only the individual invariant heterogeneities that lead to the self-selected migration and return decision, as other unobserved components not captured may lead to self-selection. Without controlling for the selection bias, it's difficult to get a proper estimator of the effect of return migration on the smoking and drinking behaviour.

Table 3.3: Baseline regressions for smoking and drinking outcomes

	Smoke	Smoke Amount	Drink	Drink Amount
Pooled OLS				
Return migration status	0.193***	2.336***	0.161***	0.380***
	(0.017)	(0.240)	(0.017)	(0.049)
Panel FE				
Return migration status	0.025***	0.340***	0.016	0.003
	(0.009)	(0.118)	(0.016)	(0.043)
N	37,754	37,754	37,754	37,754

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust Standard errors in parentheses.

3.4.1 Self-selection

In this paper, self-selection of migration is addressed, and returnees are self-selected. In the special context of rural to urban migration in China, migrants are temporary and return to the village due to those reasons were introduced in part 2. The self-selection of return is not an issue in this study. However, it's important to consider the self-selection of temporary migration.

An instrumental variable (IV) that has an effect on migration, but has no direct correlation with the smoking and drinking outcomes, is applied to control for the self-selection issue in this study.

$$Y_{i,t} = \alpha + \beta R_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \qquad (3.2)$$

where $\varepsilon_{i,t} = c_i + u_{i,t}$; and with the selection function for migration $M_{i,t}$:

$$M_{i,t} = \delta Z_{i,t} + \varphi X_{i,t} + \nu_{i,t} \qquad M_{i,t} = \begin{cases} 1, & \text{if } M^* > 0 \\ 0, & \text{otherwise} \end{cases}$$

The migration is observed only when the unobservable latent variable M^* measuring the propensity to migrate is positive. $Z_{i,t}$ is the IV, the body mass index deviation from the average. The next part introduces details of the IV.

An individual can be a return migrant only if he migrates, hence, R is only observed if this person is selected into migration ($M > 0$). The speciality of rural to urban migration in China is that all migrants return: if $M > 0$, then $R > 0$, so $R > 0 \Rightarrow M > 0$; $M > 0 \Rightarrow R > 0$. Under the specific conditions, returnees are migrants who self-selected into migration and then all migrants have to return. Controlling for "who migrates" equals to controlling for "who return". There is no issue of double selection in this study in the context of rural to urban migration in China.

It's acknowledged that the IV in this paper is more relevant for the selection of migration. This paper compares the return migrants and those current migrants based on the observables. The Table 6.11 in Appendix 6.8 shows that return migrants and current migrants are not systematically different. In fact, one can conclude that the return migrants are similar to current migrants with regards to the characteristics, such as age, marriage, education level and height. Controlling for who migrate does not exactly controls who returns. But the IV applied in this paper for the return migration would not lead to a serious bias of the result. First, return migrants are self-selected migrants with regards to observable (for instance, being taller) and unobservable characteristics. The return migrants are similar to the current migrants. And the return migration in China can be due to many reasons. For instance, not finding work, the existence of left behind

children (Démurger and Xu (2013)), finding a spouse and getting married (Fan (2003)). If one thinks of the return migration in China as negatively selected, like in international migration, this would actually enhance findings in this paper. Because return migrants smoke and drink more if they are negatively selected (which means unemployed, lower-income and so on).

The estimation of the model is based on the"$Z_{i,t}$" in the migration selection function. This paper uses the "body mass index difference from the average" to instrument migration arguing that rural to urban migrants who do manual work are positively selected based on their physiques.

3.4.2 Identification

The IV applied in this paper is "previous wave individual body mass index (height/weight) difference from average of people of the same gender, age group, province and survey time". There is abundant literature in economics on the effect of physique on wage and employment. Thomas and Strauss (1997) finds that height has a large and significant effect on wages' according to a survey data from Brazil, taller men and women earn more. Harper (2000) uses British longitudinal data to find that physical appearance has a substantial effect on earning and employment patterns, where tall men receive a pay premium, obese women experience a pay penalty. Case and Paxson (2006), using the US and the UK data, finds that height is a premium in adult earning. A more recent study by Vogl (2014) shows that taller Mexican workers are paid higher wages. He discusses that, in developing economics, return to stature derives from the greater strength of taller individuals, which leads to a productivity advantage in economics that relies on manual labour.

The literature about the effect of health on wages has also been developed (see Thomas and Strauss (1997) for a review). Sahn and Alderman (1988) finds that better health increases labour productivity as well as wage offer. Robert and Oliver (2010) finds that healthy males earn between 1.3% and 7.8% more than those in poor health.

In general, individuals with better body strength have an advantage in employment and pay, especially in manual work. Most rural to urban migrants in China are doing manual work. The China Labour Bulletin[4] shows that from 2008 to 2012, more than 80% migrants were in manual work - manufacturing, construction, wholesale and retail, transportation, catering services, household and other services. Data from Chinese National Bureau of Statistic[5] shows in 2014, there were 31.3% migrants working in manufacturing, 22.3% in construction, 11.4% in wholesale and retail, 10.2% in household and other services, 6.5% in transportation and 6.0% in catering services, in total around 90% migrants do manual jobs in 2014. It's straightforward to argue that rural to urban migrants who are doing manual work are positively selected for migration on the basis of their body status, which this paper measured by body mass index (height/weight). The measure offers a reward for being strong and in good physique.

Table 3.4 shows the mean of "height/weight difference" by return migration status for the whole sample as well as sub-sample by gender. On average the body mass of returnees is significantly above the average, while non-migrants are below average. The instrumental variable $Z_{i,t}$ is calculated as follows:

$$Z_{i,t} = (\frac{Height}{Weight})_{i,t-1} - \overline{(\frac{Height}{Weight})}_{s,a,p,t-1} \qquad (3.3)$$

Where $\overline{(\frac{Height}{Weight})}$ is the average BMI by gender (s), age groups (a), province (p) at the time of the previous wave ($t-1$). $t-1$ is actually the time of previous wave, not the previous year as there are gaps in the survey. Age groups are $[15-18], [19-60], [61-70]$. The first age group $[15-18]$ is the young adults group where human heights are still growing. The second group [19-60] is the working age group and the last group is the ageing group, when the heights of individuals shrink and weight may increase or decrease. There is sufficient variation in the body mass (height/weight) of the individual over the life cycle, even if the height is very stable at working age group but the weight varies over years. The gender group is $\{0, 1\}$ standing for male or female. The province

[4]China Labour Bulletin: http://www.clb.org.hk/en/content/migrant-workers-and-their-children

[5]Chinese National Bureau of Statistics:
$http://www.stats.gov.cn/tjsj/zxfb/201504/t20150429_797821.html$. Table 7 in Chinese

is the region where the household is located. By averaging the height/weight across various groups and generating the difference, there is enough variation over time and across individuals in different dimensions for $Z_{i,t}$ to be applied in the estimation models.

Figure 3.3 and 3.4 below show the distribution and the variation of the IV. The distribution of the IV for the non-migrants approaches the normal distribution, while that of the returnees moves right with a larger mean. The box figure shows the dispersion of the IV by gender, where 0 is for non-migrants and 1 is for returnees.

Table 3.4: IV for returnees and non-migrants

Height adjusted by weight difference	returnees	non-migrants	Diff
Whole sample	0.019	-0.053	0.072***
	(0.007)	(0.002)	(0.008)
Male	0.004	-0.061	0.065***
	(0.008)	(0.003)	(0.008)
Female	0.057	-0.045	0.102***
	(0.013)	(0.003)	(0.015)

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Standard errors in parentheses.
Year 1997-2011, 15-70 rural residents excluding migrants.

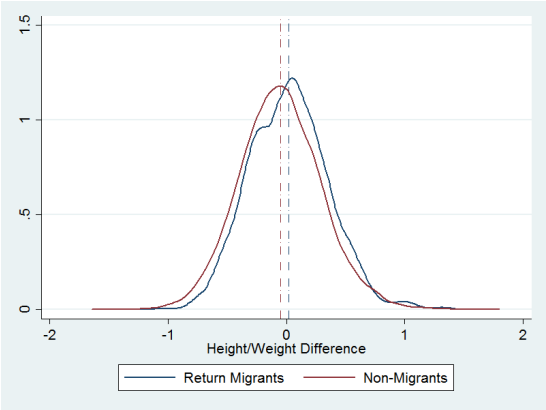


Figure 3.3: Distribution of IV by return rtatus

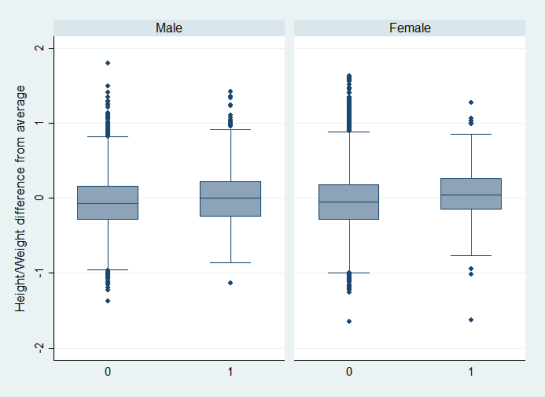


Figure 3.4: Variation in IV by return status and gender

The correlation of the IV with the selection into migration is clear. First, as men-

tioned above, the rural to urban migrants who mainly do manually work are selected by their body status. The taller and healthier rural individuals are, the more likely they are to get employed in urban areas. The IV constructed in this paper measuring the body quality gives a higher value to individuals who are taller or less obese. The IV would predict the migration and return migration. The first stage results will confirm the strong correlation.

The IV in this paper is not correlated with tobacco and alcohol consumption behaviour. The IV, which is the height/weight difference from average at previous wave $t-1$, is not correlated with the current smoking and drinking outcomes at time t. Given the year gap between the survey, waves are on average 3 years, the body quality 3 years ago is unlikely to be correlated with the current behaviour of smoking and drinking. The previous wave values of the height/weight difference would make the IV more relevant with migration as migrants are more likely to be selected for migration at wave $t-1$ if the return migration status is observed at t.

The validity of the IV rests on the assumption that historical (previous wave) body mass is not correlated with current smoking and drinking behaviour, but with the current return migration status. Accounting the difference of the height/weight from the average of peers in the same age, the gender group in the same province makes the IV more exogenous.

The IV in this paper can predict migration and is uncorrelated with the tobacco and alcohol consumption. One has strong reasons to believe that $Z_{i,t}$ is exogenous and does not impact the smoking and drinking behaviour directly. It is thus excluded from the main regression model.

3.4.3 Results

The two stage fixed effects estimation using $Z_{i,t}$ to instrument return migration status:

$$Y_{i,t} = \alpha + \beta R_{i,t} + \gamma X_{i,t} + \varepsilon_{i,t} \qquad (3.4)$$

where $R_{i,t} = \delta Z_{i,t} + \varphi X_{i,t} + \nu_{i,t}$ and $\varepsilon_{i,t} = c_i + u_{i,t}$.

$Y_{i,t}$ is the smoking and drinking measurements. $R_{i,t}$ is the indicator for return migration status. $Z_{i,t}$ instruments the return migration status. $E(u_{i,t}|Z_{i,t}, c_i) = 0, t = 1, ..., 6$. The coefficient on the IV (δ) in the first stage regression is expected to be positive and significant.

Table 3.5 presents the fixed effects and instrumental variable estimation results for smoking and drinking outcomes for the 15 to 70-year-old rural residents, excluding migrants. The model includes characteristics at the individual, household and village levels, plus time and individual fixed effects in each regression.

In this paper, the dependent variables are actually binary variable and censored outcomes. The linear regression is applied because the non-linear models can not accommodate the combined instrument variable and fixed effects estimation. Actually, the outcomes in the study are discrete choices or censored data because the nonlinear models like Probit/Logit and Tobit produce conditional expectation functions that respect limited dependent variable (Angrist and Pischke (2008)). However, the use of the individual fixed effects in the models prohibits the use of limited dependent variable maximum likelihood methods (Antman (2011a)). Furthermore, there are some advantages for using linear models for the nonlinear outcomes as stated in chapter 2. First, the linear models allow this paper to apply the individual fixed effects and the instrumental variable approach to deal with the endogeneity issues. Second, it is easy to estimate and interpret in terms of marginal effects. Third, robust clustered standard errors in this study would easily deal with the heteroskedastic problem in the linear regressions, while robust cluster standard errors may be difficult to implement in some nonlinear models. Fourth, comparing to choosing any arbitrary nonlinear model without clear mind about the conditional expectation function forms, linear models can be a preserved choice. Some may ponder why the Heckman model is chosen, instead of the IV approach. Heckman model needs a strong assumption about the joint normal distribution, which is doubted by more and more economists, while the IV approach is more robust to distributional misspecification. Others may also ask why the first difference

model is selected, instead of the fixed effects model? The answer is when time span is 2, the two types of models are the same. However, when time span is long, the fixed effects models are more efficient, Wooldridge (2010). Also if the time span is long, there may be a problem with first difference method regarding the serial correlation of the error terms. Above all, the linear model with fixed effects and the instrumental variable is applied in this analysis.

In fact, migrants who do manual jobs usually are selected on the basis of their height and weight, e.g, the strength of their body. Significantly, strength is important in the manual job sectors.

The upper part of the table presents the results of the first stage, and the lower part under "Second Stage" displays the second stage results. In each of the model for 4 outcomes, the number of observations, the R squares of the second stage regressions are also presented.

Columns under "Smoke & Smoke Amount" in Table 3.5 present the fixed effects instrumental variable (FEIV) regression results of smoke outcomes on return migration status. From the coefficients on return migration status in the second stage regressions, the positive and significant βs $(0.624, 6.293)$ suggest that return migrants are more likely to smoke, and smoke more. The change of return migration stages from 0 to 1 would lead to a 0.6 unit increase in smoking incidence and 6.3 unit increase in tobacco. From the last two columns, the results indicate that return migrants have a larger probability to drink and to drink more frequently. Hence, it can be said that the effect of return migration on drinking outcome is positive and significant $(0.733, 1.16)$.

It is important to address that the coefficients of return migration status in FEIV regressions are quite larger than the ones in the panel FE and Pooled OLS models (Table 3.3). In fact, the IV approach can be biased, especially when the instrumental variable is weak. In the case of this study, it has shown that the instrumental variable is strong, from the first stage results and the tests. "Given the large differences in estimates between FE and FEIV, one would not feel comfortable relying on the FE estimates" Wooldridge (2010). The variation in the coefficients is reasonable and shows that the IV approach

corrects the bias of the selection of migration (return) in the FE model.

Table 3.5: Smoke and drink outcomes, FEIV regressions

	Smoke	Smoke Amount	Drink	Drink Amount
First Stage				
Heigh/Weight difference	0.188***	0.188***	0.188***	0.188***
	(0.027)	(0.027)	(0.027)	(0.027)
First stage F-statistics	47.98***	47.98***	47.98***	47.98***
Prob $>$ F	(0.000)	(0.000)	(0.000)	(0.000)
Second Stage				
Return migration status	0.624***	6.293***	0.733***	1.160**
	(0.141)	(1.579)	(0.198)	(0.483)
Overidentification test	exactly	exactly	exactly	exactly
	Identified	Identified	Identified	Identified
Underidentification test	35.48***	35.48***	35.48***	35.48***
Weak identification test	47.98***	47.98***	47.98***	47.98***
Weak instrument robust inference	32.02***	21.22***	14.96***	8.01***
Durbin-Wu-Hausman test	20.29***	15.79***	12.19***	5.36**
Individual Controls	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes
Community Controls	Yes	Yes	Yes	Yes
Number of fixed effects	9,874	9,874	9,874	9,874
R square	0.900	0.884	0.625	0.699
N	37,754	37,754	37,754	37,754

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Cluster robust standard errors in parentheses.

On an average, return migrants would have a 60% to 70% higher probability to smoke and drink, as compared to the non-migrants, and the increase in the amount of smoking and drinking with a change in return migration status is large. The return migrants would consume 6.3 unit more tobacco and 1.2 unit more alcohol. The results support the hypothesis that return migrants are more likely to smoke and drink and smoke and

drink more than non-migrants, and the effect is strong and statistically significant.

The relevance of the IV is supported from the first stage regressions. The coefficients of $Z_{i,t}$ in the first stage regression are positive and significant (0.188). This suggests the IV is relevant. Return migrants are migrants who are positively selected based on their body status. The argument on the positive selection into migration and return can also be supported (see Table 6.7 in Appendix 6.4) in various models. The F tests on excluded instrument are large (about 48) in all the four regression models. The F statistics are larger than the conventional number 10. There is one IV for the one endogenous variable, return migration status. The first stage results mean there is no need to worry about the weak instrument issue. The set of tests on identification and weak instrument show the IV works well.

As the instrumental variable methods are applied in these regressions, the test of endogeneity of the endogenous variable has been implemented and reported in the regression tables. The test implemented is the Durbin-Wu-Hausman test (Durbin (1954), Hausman (1978),Wu (1974)). The "Durbin-Wu-Hausman test of endogeneity" reports the Durbin-Wu-Hausman Chi Square test. The results show that the return migration status should be treated as an endogenous variable.

The FEIV estimation suggests that return migration has a positive effect on tobacco and alcohol consumption behaviour. Return migrants are more likely to smoke and drink and consume more tobacco and alcohol than the non-migrants. Next, this paper studies the smoking and drinking behaviour of the male migrants. Female return migrants are excluded as they are less in number and only a few females smoke or drink. The FEIV regression results for male sub-samples are presented in Table 3.6.

There are 4 models in Table 3.6 that present the results of male sub-sample. The first stages of the regression are presented in the upper part of the table followed by the second stage results.

The first stage results show that the IV is strong. The F test on excluded instruments for all the regressions in Table 3.6 are large (24). The coefficients of IV in the first stage regression are positive and significant (0.18). The IV applied is strongly correlated with

	Male			
	Smoke	Smoke Amount	Drink	Drink Amount
First Stage				
Heigh/Weight difference	0.177***	0.177***	0.177***	0.177***
	(0.036)	(0.036)	(0.036)	(0.036)
First stage F-statistics	24.36***	24.36***	24.36***	24.36***
Prob > F	(0.000)	(0.000)	(0.000)	(0.000)
Second Stage				
Return migration status	0.756***	6.194**	0.805**	1.538*
	(0.254)	(2.651)	(0.342)	(0.848)
Overidentification test	exactly	exactly	exactly	exactly
	Identified	Identified	Identified	Identified
Underidentification test	19.56***	19.56***	19.56***	19.56***
Weak identification test	24.36***	24.36***	24.36***	24.36***
Weak instrument robust inference	16.80***	7.31***	5.71**	4.56**
Durbin-Wu-Hausman test	10.46***	5.29**	4.97**	3.95**
Individual Controls	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes
Community Controls	Yes	Yes	Yes	Yes
Number of fixed effects	4,939	4,939	4,939	4,939
R square	0.800	0.827	0.520	0.588
N	18,689	18,689	18,689	18,689

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust cluster standard errors in parentheses.

Table 3.6: Smoke and drink outcomes for males, FEIV regressions

the migration status and can predict migration and return migration.

The coefficients on return migration in the second stage are positive and significant for all the four outcomes, except the drinking amount. Male return migrants have a higher incidence of drinking and smoking, and they also consume more tobacco. The coefficients on smoking and drinking outcomes for the male sample are larger than the coefficients in Table 3.5, except for the smoking extent. This implies that the overall significant effect of return migration on smoking and drinking behaviour is driven by males.

The return migration has an impact on the males' behaviour. The effect of return migration on consumption in tobacco and alcohol is driven by males. For males alone, the effect of return migration on smoking and drinking consumption behaviour is strong. The male returnees would have around 80% higher incidence to smoke and drink and smoke 6 unit more. This paper provides evidence that male return migrants are more likely to smoke and drink, and would consume more tobacco. The pooled OLS and linear fixed effects models have underestimated the impact of return migration on smoking and drinking.

3.5 Difference in Difference

In the previous sections, this paper focused on the causal effect of return migration on the smoking and drinking behaviour, comparing the smoking and drinking behaviour of returnees and non-migrants. Return migrants smoke and drink more than non-migrants and the impact is significant and positive only for males.

It's also important to know the mechanism behind the finding. Why do male return migrants smoke and drink more? Do they experience significant changes in smoking and drinking behaviour before and after migration? To investigate these issues, this paper further evaluates the smoking and drinking patterns of males, before and after migration to check the behaviour change. Rural males excluding current migrants are

divided into two groups, those who have migration experience are the treatment group, while those who never migrated are in the control group.

3.5.1 Pooled DiD

Firstly, as the treatment (migration) may happen at any time of the survey, the treatment group and control group are pooled together, regardless of the timing of treatment. The intuitive way to show the difference between the two groups is to present the average outcomes over the years by the control and treatment group. The 4 figures (3.5-3.8) below show the average smoking and drinking outcomes by treatment and control groups over the years.

These figures show that, for all the four types of drinking and smoking measures, there is a significant difference between the control group and the treatment group over the years. Overall, the treatment group is more likely to smoke and drink, and smoke and drink more frequently than the control group for all the survey years. This shows that the treatment group smokes and drinks more, even in the pooled setting, where the exact timing of the treatment is not clear.

The figures also show the "parallel trend" for the treatment and control group at the early waves. For example, there are decreasing trends in the control and treatment group in the years 1997 to 2000 for drinking, while there is an increasing parallel trend for drinking amount measure. There are both decreasing trends from 1997 to 2004 for the control and treatment group for the smoking and smoking amount measures. This parallel trend is observed in the early years, which are very likely to be the "before migration" years. These trends show important information for the difference in difference (DiD) setting because, for comparison in DiD, the control and treatment groups should have similar "before treatment" patterns. These figures provide evidence for the parallel trend assumption needed for the DiD setting.

Table 3.7 shows the DiD regression results of the pooled DiD setting. This pseudo setting verifies the control and treatment group, but the timing of treatment is not ver-

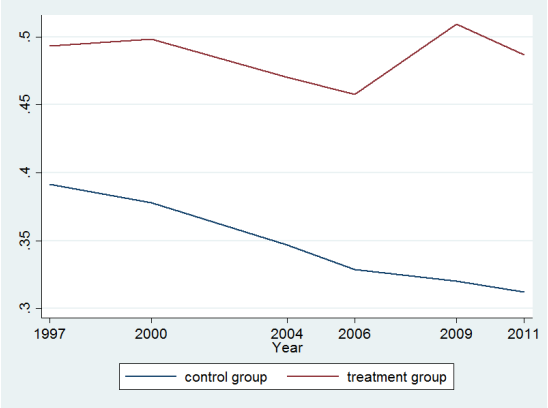


Figure 3.5: Smoke by treatment and control group, rural males

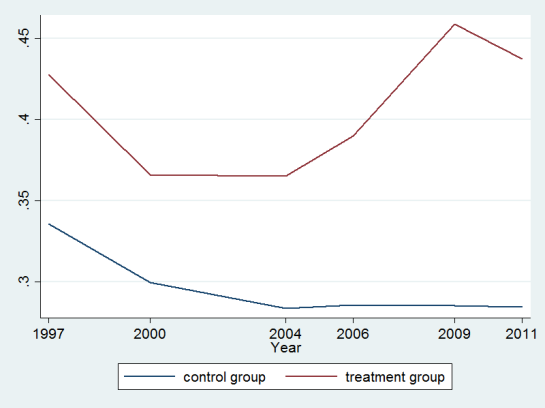


Figure 3.6: Drink by treatment and control group, rural males

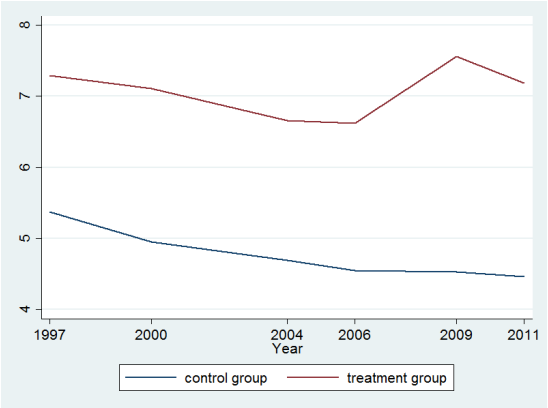


Figure 3.7: Smoke amount by treatment and control group, rural males

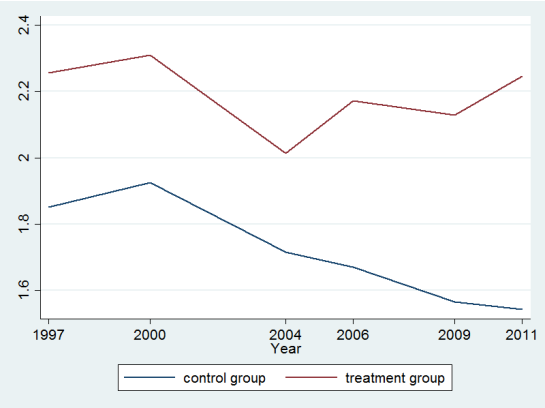


Figure 3.8: Drink amount by treatment and control group, rural males

ified. The results show that the rural male migrants have a significant increase in the smoking and drinking incidences and amount.

Another pseudo setting treats each wave of the survey as the timing of treatment. The treatment and control groups are the same as the previous setting but assume that migration happens in the year 2000, 2004, 2006 or 2009. Table 3.8 shows the DiD regression results in this pseudo setting. The column "diff-in-diff" shows the before and after "migration" change in the smoking and drinking behaviour between the control and

treatment group. The results suggest that there are positive and significant increases in smoking and drinking for the rural male residents who have migration experience.

Table 3.7: Pooled difference in difference: overall

	Difference baseline	Difference follow up	diff-in-diff
smoking	0.083***	0.099***	0.017***
	(0.018)	(0.014)	(0.005)
drinking	0.045***	0.065***	0.019***
	(0.018)	(0.014)	(0.004)
smoke amount	1.644***	1.843***	0.200***
	(0.296)	(0.232)	(0.074)
drink amount	0.246*	0.313***	0.067**
	(0.127)	(0.099)	(0.031)

Rural male residents, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Standard errors in parentheses.

Table 3.8: Pooled difference in difference: single year treatment

		Difference baseline	Difference follow up	diff-in-diff
2000	smoking	-0.077***	0.015	0.092***
		(0.016)	(0.011)	(0.020)
	drinking	-0.076***	0.019	0.095***
		(0.017)	(0.012)	(0.021)
	smoke_Amt	-0.715**	0.263	0.978***
		(0.300)	(0.209)	(0.365)
	drink_Amt	-0.723***	-0.354***	0.369**
		(0.142)	(0.099)	(0.173)
2004	smoking	-0.060***	0.028**	0.088***
		(0.013)	(0.013)	(0.019)
	drinking	-0.065***	0.036***	0.101***
		(0.015)	(0.014)	(0.020)
	smoke_Amt	-0.604**	0.463*	1.067***
		(0.249)	(0.237)	(0.344)
	drink_Amt	-0.701***	-0.253**	0.449***
		(0.118)	(0.112)	(0.163)
2006	smoking	-0.042***	0.036**	0.078***
		(0.012)	(0.015)	(0.019)
	drinking	-0.047***	0.044***	0.091***
		(0.013)	(0.016)	(0.021)
	smoke_Amt	-0.399*	0.552**	0.950***
		(0.218)	(0.279)	(0.354)
	drink_Amt	-0.589***	-0.233*	0.357**
		(0.103)	(0.016)	(0.168)
2009	smoking	-0.025**	0.034	0.059**
		(0.010)	(0.021)	(0.024)
	drinking	-0.023**	0.032	0.055**
		(0.011)	(0.023)	(0.026)
	smoke_Amt	-0.163	0.441	0.604
		(0.191)	(0.395)	(0.438)
	drink_Amt	-0.548***	-0.123	0.425**
		(0.090)	(0.187)	(0.208)

Rural male residents, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Standard errors in parentheses.

3.5.2 DiD in Separate Panels

The two cases of the pooled DiD show the validity of the DiD setting by presenting the parallel trend and significant increase in smoking and drinking for rural males who have migration experience. 9 panels are created by the real migration timing. In each panel, there are before migration and after migration groups, as well as control and treatment groups. Table 3.9 shows the construction of the panels. Migration year shows at which year (or years) migration occured. The period before the timing of migration is before treatment and the period after the migration is the after treatment group. For instance, Panel 1 is constructed as follows: Taking individuals who migrate at 2000 (not migrate at any other years) as the treatment group. Individuals who never migrate are in the control group. Time before 2000 is the before migration group and time after 2000 is the after migration group. The same rationale applies to panel 2 to panel 4. Panel 5 takes individuals who migrate at year 2000 and 2004 (not migrate at all other years) as the treatment group and individuals who never migrate as control group. The years before 2000 are before migration group and years after 2004 are after migration group. Panels 6 and 7 are constructed accordingly. The panel 8 takes individuals who migrate at year 2000, 2004 and 2006 as treatment group and those never migrate individuals as control group. The year before 2000 is before migration group and the year after 2006 is after migration group.

Table 3.10 presents descriptive statistics of before and after migration smoking and drinking patterns for rural male residents, who are in the treatment group in each panel.

Table 3.9: Construction of difference in difference panels

migration year	2000	2004	2006	2009
	panel 1	panel 2	panel 3	panel 4
migration years	2000 & 2004	2004 & 2006	2006 & 2009	
	panel 5	panel 6	panel 7	
migration years	2000, 2004 & 2006	2000, 2006 & 2009		
	panel 8	panel 9		

CHNS, 1997-2011 15-70 rural male residents.

Table 3.10: Comparison of before and after migration behaviour

	treatment only	before migration	after migration
panel 1	proportion of smokers	45.34%	63.64%
	proportion of drinkers	45.78%	61.63%
panel 2	proportion of smokers	58.06%	70.29%
	proportion of drinkers	50.21%	64.81%
panel 3	proportion of smokers	56.96%	68.49%
	proportion of drinkers	48.45%	62.45%
panel 4	proportion of smokers	57.18%	75.44%
	proportion of drinkers	50.39%	61.69%
panel 5	proportion of smokers	51.43%	90.48%
	proportion of drinkers	37.14%	63.93%
panel 6	proportion of smokers	70.65%	68.12%
	proportion of drinkers	52.50%	66.15%
panel 7	proportion of smokers	68.67%	90.91%
	proportion of drinkers	47.835	80.95%
panel 8	proportion of smokers	21.43%	52.94%
	proportion of drinkers	35.71%	66.67%
panel 9	proportion of smokers	52.54%	93.33%
	proportion of drinkers	38.89%	53.33%

15-70 years old rural males with migration experience.

The before and after migration changes in smoking and drinking behaviour are obvious. For all panels, the proportions of smokers and drinkers for rural men in the treatment group increases after migration. The increases are larger for panels 5 to 9, where rural men migrate in more than one wave. This may be a simple evidence proving that migration has an influence on the rural males, as the longer the migration, the larger the change in smoking and drinking behaviour. The migration and urban living experience change the smoking and drinking behaviour. However, the smoking and drinking behaviour of non-migrants may also show a similar change. There is a need to compare this behaviour change of migrants with that of the non-migrants in the DiD setting.

Let t be the timing of migration. It can be migration year (years) stated in Table 3.9. Let d_t be the dummy variable that switches on for time after t, and M_i be a dummy for rural male residents who have migration experience at time t (not other years). β_1 to β_4 are the parameters, $\varepsilon_{i,t}$ is the error term. So, the basic regression DiD is modelled as follows:

$$y_{i,t} = \beta_1 + \beta_2 M_i + \beta_3 d_t + \beta_4 (M * d)_{i,t} + \varepsilon_{i,t} \qquad (3.5)$$

To capture the treatment effect of migration on the treatment group, the effect is estimated for each of the 9 panels. The results are presented in Table 3.11.

The "Difference baseline" column shows the difference in smoking and drinking outcomes before t and these differences are significant. The negative and significant coefficients suggest that the migrants are less likely to smoke and drink, and in lesser frequency than non-migrants before the migration. Potential migrants smoke and drink less before migration, then after migration, they alter their behaviour. This implies that the finding using fixed effects and instrumental variable strategy is robust and return migration has a positive effect on the drinking and smoking consumption of males.

The column "Difference follow up" shows the after migration difference of smoking and drinking behaviour between the treatment and control groups. Most of the after migration differences are positive, which means the treatment group smokes and drinks more after migration. The column "diff-in-diff" shows the DiD estimation results for β_4. All the DiD coefficients are positive and most of them are significant, except some that reflect the drink amount outcome.

The effect of migration on smoking incidence is significant, except in panel 6. The effect is around 0.1 to 0.2 in the first 4 panels. However, in panel 5 to panel 9, when t are in multiple waves, the number can be higher than 0.5. Similarly, the effect of migration on drinking incidence is positive and significant except in panel 9. The magnitude of the impact is about 0.1 to 0.2 in the first four panels, while the number can be as high as 0.4 in the last four panels.

Table 3.11: Difference in difference estimation: basic setting

		Difference baseline	Difference follow up	diff-in-diff (β_4)
panel 1	smoking	-0.336***	-0.101***	0.235***
		(0.027)	(0.021)	(0.034)
	drinking	-0.190***	0.011	0.202***
		(0.030)	(0.027)	(0.039)
	smoke_Amt	-5.588***	-1.735***	3.852***
		(0.551)	(0.452)	(0.713)
	drink_Amt	-1.524***	-0.516**	1.009***
		(0.264)	(0.214)	(0.340)
panel 2	smoking	-0.209***	-0.016	0.193***
		(0.017)	(0.019)	(0.026)
	drinking	-0.141***	0.044*	0.185***
		(0.020)	(0.022)	(0.030)
	smoke_Amt	-3.417***	-0.315	3.102***
		(0.372)	(0.426)	(0.566)
	drink_Amt	-1.319***	-0.194	1.125***
		(0.175)	(0.193)	(0.260)
panel 3	smoking	-0.205***	-0.024	0.181***
		(0.015)	(0.020)	(0.025)
	drinking	-0.145***	0.022	0.167***
		(0.017)	(0.023)	(0.029)
	smoke_Amt	-3.485***	-0.317	3.168***
		(0.313)	(0.433)	(0.534)
	drink_Amt	-1.301***	-0.256	1.045***
		(0.148)	(0.198)	(0.247)
panel 4	smoking	-0.184***	0.052	0.236***
		(0.015)	(0.036)	(0.039)
	drinking	-0.126***	0.014	0.140***
		(0.017)	(0.041)	(0.044)
	smoke_Amt	-3.417***	0.853	4.270***
		(0.320)	(0.763)	(0.827)
	drink_Amt	-1.162***	-0.490	0.672*
		(0.150)	(0.354)	(0.384)

table continues next page

		Difference baseline	Difference follow up	diff-in-diff (β_4)
panel 5	smoking	-0.283***	0.175***	0.459***
		(0.074)	(0.055)	(0.092)
	drinking	-0.282***	0.035	0.317***
		(0.083)	(0.063)	(0.104)
	smoke_amt	-4.646***	3.688***	8.334***
		(1.528)	(1.315)	(2.016)
	drink_amt	-1.530**	-0.625	0.905
		(0.724)	(0.544)	(0.906)
panel 6	smoking	-0.094**	-0.036	0.058
		(0.045)	(0.052)	(0.069)
	drinking	-0.125**	0.057	0.182**
		(0.055)	(0.061)	(0.082)
	smoke_amt	-2.349**	-0.045	2.305
		(1.028)	(1.158)	(1.549)
	drink_amt	-1.072**	-0.490	0.583
		(0.471)	(0.529)	(0.708)
panel 7	smoking	-0.087*	0.210**	0.297***
		(0.048)	(0.092)	(0.104)
	drinking	-0.159***	0.210**	0.368***
		(0.059)	(0.106)	(0.121)
	smoke_amt	-2.419**	5.442***	7.861***
		(1.086)	(2.003)	(2.279)
	drink_amt	-1.666***	0.788	2.455**
		(0.503)	(0.932)	(1.059)
panel 8	smoking	-0.593***	-0.199*	0.394**
		(0.115)	(0.104)	(0.155)
	drinking	-0.298**	0.061	0.360***
		(0.130)	(0.125)	(0.181)
	smoke_amt	-8.916***	-3.660	5.257
		(2.323)	(2.404)	(3.343)
	drink_amt	-2.382**	0.528	2.911*
		(1.101)	(1.139)	(1.584)
panel 9	smoking	-0.273***	0.229**	0.501***
		(0.056)	(0.110)	(0.123)
	drinking	-0.271***	-0.064	0.207
		(0.066)	(0.125)	(0.141)
	smoke_amt	-4.096***	3.094***	8.001***
		(1.214)	(2.411)	(2.700)
	drink_amt	-2.413***	-0.224	2.189*
		(0.565)	(1.102)	(1.238)

Rural male residents, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Standard errors in parentheses.

The DiD results for smoke amount are significant and positive for all panels, except panels 6 and 8. In the first four panels, when migration timings are of single years, the treatment group smokes 3 to 4 units more, while in panels 5 to 9, they smoke about 8 units of cigarettes more. The impact of migration on drink amount is not very significant in the basic DiD model. This is consistent with the not significant results for the drinking amount of rural males in Table 3.6.

Migration has a significant impact on the smoking and drinking behaviour of rural males. The longer the migration experience, the stronger the change is. Those who have migration experience change their smoking and drinking behaviour. Before migration, they smoke and drink less than the non-migrants. After migration, their tendency to smoke and drink increases. This result is consistent with the findings from the instrumental variable strategy when controlling for the self-selection of migration.

However, it is often argued that the simple DiD estimation may be biased because "treatment" individuals are people who are self-selected into migration. The "treatment" group is not exogenous. As in the DiD, the treatment is always associated with a policy change or an exogenous shock, and DiD is the method to study the impact of the policy or shock on the treatment group. Migration is self-selected and treating migration as an exogenous shock would lead to a biased result.

To deal with this potential bias, this paper combines the IV strategy within the DiD setting. Instead of estimating equation 3.5 directly, a two-stage estimation is applied. The first stage predicts the treatment dummy by instrumental variable, while the second stage estimates the DiD.

$$y_{i,t} = \beta_1 + \beta_2 M_i + \beta_3 d_t + \beta_4 (M * d)_{i,t} + \beta_5 X_{i,t} + \varepsilon_{i,t} \qquad (3.6)$$

Now $M_{i,t} = \delta Z_{i,t} + \varphi X_{i,t} + \nu_{i,t}$ which is the first stage regression. $Z_{i,t}$ is the instrumental variable "body mass index difference from average at the previous wave". $X_{i,t}$ is a set of controls, such as individual age, marriage status, education level and so on.

First stage regresses $M_{i,t}$ on instrumental variable, and predict M_i for each individual. The second stage runs the DiD regression with controls using the predicted M_i

as the treatment dummy. It's noticed that predicted M_i from the first stage regression is not actually a treatment dummy, while predicted values of $M_{i,t}$ from the first stage regression are used as the treatment dummy.

The regression results of "DiD + IV" strategy are presented in Table 3.12. Column "Difference baseline" shows the before migration differences between control and treatment groups for smoking and drinking outcomes across the 9 panels. Column "Difference follow up" presents the after migration differences and the column "diff-in-diff" is the DiD + IV results from the second stage estimation, which is β_4 in equation 3.6.

For the panels, the differences in smoking incidence before migration varies. Some are positive and significant, while some are non-significant. However, for all the panels, the differences in smoking incidence after migration are all positive and significant. This shows that migrants are more likely to smoke upon return, which supports our finding in FEIV estimations. For all panels, the DiD estimation results for smoking incidence are positively significant, suggesting the strength of the effect of migration experience on the smoking behaviour of rural males.

Behaviour change in drinking incidence before and after migration is also observed. Before migration, the differences of drinking incidences between control and treatment groups are either significantly negative or non-significant. After migration, there is a large shift in the drinking incidence, thus there are no differences in drinking likelihood between the control and treatment groups after migration. This shift leads to significant DiD results for drinking incidence in all the panels, except panel 7.

Both before and after migration differences in smoke amount are positively significant in all the panels. The treatment group smoke more than the control group before migration and after migration. The DiD results show that migration has a strong positive effect on the smoke amount. Migrants increase their smoking amounts when they return even if they smoke more than the non-migrants before migration. The coefficients for drinking amount before migration are negative or non-significant. Furthermore, those who experienced migration catches up or exceeds the non-migrants in terms of drinking amount. This change makes the DiD results significantly positive.

Table 3.12: DiD and IV estimation results

		Difference baseline	Difference follow up	diff-in-diff
panel 1	smoking	-0.397	1.095***	1.492***
		(0.268)	(0.225)	(0.194)
	drinking	-1.458***	-0.038	1.420***
		(0.299)	(0.251)	(0.218)
	smoke_Amt	16.627***	45.270***	28.643***
		(5.798)	(4.876)	(4.190)
	drink_Amt	-8.350***	0.933	9.283***
		(2.636)	(2.197)	(1.918)
panel 2	smoking	-0.031	1.833***	1.535***
		(0.223)	(0.166)	(0.163)
	drinking	-0.833***	0.204	1.038***
		(0.256)	(0.245)	(0.186)
	smoke_Amt	13.355***	42.696***	29.341***
		(4.912)	(4.749)	(3.532)
	drink_Amt	-2.541	3.375	5.916***
		(2.216)	(2.121)	(1.607)
panel 3	smoking	0.623***	1.673***	1.050***
		(0.215)	(0.228)	(0.157)
	drinking	-0.240	0.544**	0.784***
		(0.246)	(0.258)	(0.178)
	smoke_Amt	24.459***	42.357***	17.898***
		(4.685)	(4.935)	(3.337)
	drink_Amt	1.982	6.102***	4.120***
		(2.112)	(2.212)	(1.524)
panel 4	smoking	0.716***	1.598***	0.882***
		(0.226)	(0.276)	(0.207)
	drinking	-0.044	0.582*	0.626***
		(0.259)	(0.312)	(0.233)
	smoke_Amt	27.977***	43.877***	15.900***
		(4.985)	(5.998)	(4.416)
	drink_Amt	2.971	7.201***	4.482**
		(2.242)	(2.696)	(2.009)

table continues next page

		Difference baseline	Difference follow up	diff-in-diff
panel 5	smoking	0.046	1.324***	1.277***
		(0.325)	(0.268)	(0.257)
	drinking	-1.461***	-0.251	1.211***
		(0.365)	(0.301)	(0.289)
	smoke_Amt	23.132***	49.788***	26.655***
		(7.156)	(5.882)	(5.641)
	drink_Amt	-9.057***	0.010	9.067***
		(3.207)	(2.623)	(2.534)
panel 6	smoking	0.588**	1.609***	1.021***
		(0.278)	(0.275)	(0.234)
	drinking	-0.433	0.260	0.692**
		(0.325)	(0.315)	(0.271)
	smoke_Amt	17.368***	41.096***	23.727***
		(6.267)	(6.102)	(5.238)
	drink_Amt	-0.642	5.028*	5.670**
		(2.820)	(2.729)	(2.347)
panel 7	smoking	0.701***	1.297***	0.596**
		(0.264)	(0.319)	(0.255)
	drinking	-0.036	0.484	0.520*
		(0.308)	(0.365)	(0.291)
	smoke_Amt	22.001***	34.831***	12.829**
		(5.959)	(7.057)	(5.579)
	drink_Amt	1.774	7.022**	5.248**
		(2.686)	(3.174)	(2.529)
panel 8	smoking	0.316	1.465***	1.149***
		(0.394)	(0.334)	(0.330)
	drinking	-1.412***	-0.254	1.158***
		(0.446)	(0.377)	(0.373)
	smoke_Amt	23.051***	47.507***	24.456***
		(8.665)	(7.294)	(7.261)
	drink_Amt	-10.874***	0.087	10.961***
		(3.876)	(3.242)	(3.239)
panel 9	smoking	0.304	1.588***	1.284***
		(0.321)	(0.364)	(0.317)
	drinking	-0.748**	-0.008	0.740**
		(0.378)	(0.420)	(0.365)
	smoke_Amt	11.069	35.684***	24.615***
		(7.335)	(8.170)	(7.121)
	drink_Amt	-3.210	5.897	9.107***
		(3.277)	(3.618)	(3.135)

Rural male residents, $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Standard errors in parentheses.

Overall, DiD + IV estimation results suggest that migration has a positive and significant impact on the smoking and drinking behaviour of male migrants. Migration experience increases the likeliness of male migrants to drink and smoke, and in higher amount. From FEIV results, male return migrants are more likely to smoke and drink and smoke more than non-migrants. From DiD models in the pooled settings, or separated panels settings comparing the before and after migration changes, the evidence suggests that rural males increase their smoking and drinking incidences and frequency after migration.

Empirical evidence from this study suggests that migration experience has changed the smoking and drinking behaviour of rural males.

3.6 Robustness Checks

The findings in this paper indicate that migration changes the smoking and drinking pattern of rural males. It has been argued that, simply because of their higher income and savings, migrants will consume more alcohol and tobacco, as they have accumulated more capital.

Smoking and drinking consumption of male return migrants may be influenced by income. It's important to check how the income impacts the smoking and drinking consumption. Income measures, such as household and individual income affects the migration and return decision, as well as the tobacco and alcohol consumption, are excluded in the main regressions. This exercise controls the income measures to show that the results are consistent.

Household income and individual income for rural male residents are both measured in nature logarithm form. The incomes are valued in 2011 money for different waves of the survey. Equation 3.4 is estimated with the FEIV strategy for the rural male sample with the income variables as an extra control. The results are presented in Table 3.13.

The four models under "Household income" are the models with household income

as additional control and the four models under "Individual income" are models with individual income as extra controls. The first stage shows the results of the first stage regressions, while the second stage shows the results of the second stage regressions as well as the number of observations, the R square of the second stage. All the models include the individual, household and village controls together with individual and time fixed effects.

The first stage presents the good results for the IV. The IV has a positive and significant effect on the return migration status (0.18), while the first stage F statistics are about 24. The second stage regression results indicate that the coefficients on return status are large and significant except for the drinking outcome. Adding the household income or individual income as extra controls does not impact the results. The coefficients on return migration in this robustness check, where household income or individual income are included are consistent with results in the main finding in Table 3.5 and the results for the male sample in Table 3.6.

As for the impact of income on the smoking and drinking consumption, the coefficients of the income measures in the second stage indicate that household income has no significant effect on the smoking pattern of males, but have a positive effect on drinking incidence and drinking amount. The individual income has a significant effect on both the smoking and drinking outcomes. The coefficients of individual income measures are all significant and large in magnitude as compared to the household income measure. These results are plausible and expected. Income is expected to have a positive impact on the consumption of tobacco and alcohol. The individual income of the males, which is more directly controlled by males themselves, would be more closely related to their own consumption.

This exercise provides evidence that income, especially individual income, has a positive effect on the smoking and drinking behaviour. The impact of return migration on smoking and drinking patterns of males is strong. After income measures are controlled, there is still strong and consistent evidence to indicate that male migrants are more likely to smoke and drink, and smoke and drink more than the non-migrants.

	Household income				Individual income			
	Smoke	Smoke_Amt	Drink	Drink_Amt	Smoke	Smoke_Amt	Drink	Drink_Amt
First Stage								
Heigh/Weight difference	0.177***	0.177***	0.177***	0.177***	0.177***	0.177***	0.177***	0.177***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
First stage F-statistics	24.40***	24.40***	24.40***	24.40***	24.38***	24.38***	24.38***	24.38***
Prob > F	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Second Stage								
Return migration status	0.757***	6.216**	0.800**	1.525*	0.746***	6.048**	0.763**	1.476*
	(0.254)	(2.658)	(0.340)	(0.843)	(0.254)	(2.635)	(0.338)	(0.848)
Income measurement	-0.001	-0.023	0.005**	0.012**	0.002***	0.030***	0.009***	0.038***
	(0.001)	(0.015)	(0.002)	(0.006)	(0.001)	(0.009)	(0.001)	(0.009)
Overidentification test	exactly	exactly	exactly	exactly	exactly	exactly	exactly	exactly
	Identified	Identified	Identified	Identified	Identified	Identified	Identified	Identified
Underidentification test	19.59***	19.59***	19.59***	19.59***	19.63***	19.63***	19.63***	19.63***
Weak identification test	24.40***	24.40***	24.40***	24.40***	24.38***	24.38***	24.38***	24.38***
Weak instrument robust inference	16.88***	7.32***	5.71**	4.55**	16.69***	7.10***	5.34**	4.26**
Durbin-Wu-Hausman test	10.49***	5.29**	4.97**	3.95**	10.19***	5.10**	4.59**	3.73*
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Community Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of fixed effects	4,939	4,939	4,939	4,939	4,939	4,939	4,939	4,939
R square	0.800	0.827	0.521	0.589	0.802	0.828	0.528	0.591
N	18,689	18,689	18,689	18,689	18,689	18,689	18,689	18,689

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Robust cluster standard errors in parentheses.

Table 3.13: Robustness check: income in male migrants smoking and drinking behaviour

3.7 Conclusion

This is the first empirical work that examines the consumption of return migrants. This paper focuses on the consumption of tobacco and alcohol of Chinese return migrants. The study attempts to seek answers to the questions like, are return migrants more likely to smoke; and, do they smoke more and drink more than non migrants? This paper uses a national panel survey and applies two empirical strategies to answer these questions and quantify the effect of migration on alcohol and tobacco consumption.

By dealing with the selection of migration and comparing the before and after migration behaviour change, this study provides empirical evidence that the male migrants are more likely to smoke and drink, and smoke more than male non-migrants. Migration experience has a positive and significant effect on the drinking and smoking behaviour of rural males.

Polled DiD settings, a basic DiD model and the DiD + IV model provides strong evidence that there are significant before and after migration changes in smoking and drinking behaviours. Rural males who have experienced migration will increase their likelihood to smoke and drink, and its amount.

It's not simply the income effect of migration that alters the behaviour of rural males. The household income has no effect on the smoking, but it positively affects the drinking of male returnees. Moreover, the individual income increases the smoking and drinking of rural males. However, the impact of return migration is consistent after controlling the income variables. This suggests that the smoking and drinking pattern change of male returnees is simply not a result of the income effect.

This paper does not provide direct evidence of the mechanism by which the migration and return migration impacts the smoking and drinking patterns of male returnees. However, from the existing literature on international migration, the results in this paper suggest that male migrants adapt to the urban lifestyle, and are exposed to the urban norms. They cultivate habits and increase smoking and drinking probabilities and amounts during migration. Upon their return to their villages, they bring back their

smoking and drinking behaviours.

In order to better illustrate the "transfer of social norms" and "adapt to the behaviours from migration destination", this paper obtains the urban data and compares the smoking and drinking pattern of males (as mentioned previously, females are not relevant in the analysis). The Table 6.10 in Appendix 6.7 illustrates the mean of smoking and drinking probabilities and amount, separately for urban and rural males. The differences and the t-tests show that urban males do smoke and drink more than rural males. This can be an evidence that rural migrants adapt to the behaviour from the destination and bring back this behaviour upon return. It is argued that urban residents are more educated, on an average, and hence, smoke and drink less. However, education level has no significant impact on the health knowledge (as the study in Chapter 4 indicates).

The findings of this paper contribute to the growing literature on the impact of return migrants, especially the emerging empirical works on how migration transfers social norms. For policy makers, this may be critical for understanding the fact that return migrants consume more alcohol and tobacco. Also, public policy should address the health and social problems of smoking and drinking. It is also important for policy makers for reminding migrants to improve their financial literacy, entrepreneurship, and planning for the future, instead of consuming more tobacco and alcohol.

4

Migration, Mass Media and Diet Knowledge

4.1 Introduction

Healthy diet knowledge requires more public and academic attention. According to the World Health Organization (WHO), "unhealthy diet and lack of physical activity are leading global risks to health". Diet knowledge affects diet behaviour and hence, the health outcome of the population, which has "long-term effects on the economic development" (Strauss and Thomas (1998)).

The diet knowledge changes over time. WHO states that [1]"the increased production of processed food, rapid urbanization and changing lifestyles have led to a shift in dietary patterns. People are now consuming more foods high in energy, fats, free sugars or salt/sodium; and many do not eat enough fruit, vegetables and dietary fibre such as

[1]WHO Healthy Diet Fact Sheet: http://www.who.int/mediacentre/factsheets/fs394/en/

whole grains."

Diet behaviour depends on numerous factors like the preferences of the individual, the cost of food, income and the knowledge of diet. In economic literature, much focus has been given to preferences and economic factors, such as education, cost and income (Kenkel (1991), Handbury et al. (2015)). However, as stated in Blaylock et al. (1999), "Information and knowledge are given prominent play as counter balances to the economic forces driving eating habits". Therefore, diet knowledge must play an important role.

This paper studies the diet knowledge in China. Considering the two vital channels of information transmission, migration and mass media exposure, the research questions asked are, what are the determinants of diet knowledge of rural residents in China; and do village migration and/or mass media exposure affect the knowledge of villagers? To answer these questions, a panel survey, China Health and Nutrition survey, is analysed. The regression models include individual fixed effects, time fixed effects and a set of control variables of individual, household and village characteristics. The results show that both mass media and migration density significantly influence the diet knowledge of rural residents. There are strong and positive externalities of village migration, regarding the diet knowledge of individuals. This indicates that the migrants bring the health diet knowledge from the migration destination to villages and this knowledge, in turn, spreads to other residents of the area. At the same time, mass media also transmits the knowledge from the outside world to the rural residents.

Diet impacts the physical health of people. As stated by Willett (1994), "many studies have implicated dietary factors in the cause and prevention of important diseases." Diet knowledge can be determined by factors such as education, income levels, the culture and social norms. Other issues may also be important, for instance, mass media. Zimmerman FJ (2014) finds that "televised food advertising has a strong effect on individual food choice". Exposure to mass media affects the diet knowledge of the targeted population, but empirical evidence about the impact of the media on diet knowledge is rare in the economics literature.

Studies in the economics of migration suggest that migration experience can transfer the social norms and beliefs. For instance, Bertoli and Marchetta (2015) finds that Egyptian return migrants adjust their fertility choices to the norms in migration destination countries. However, there is a lack of evidence regarding how migration experience alters the knowledge of people, especially the diet knowledge.

Mass media affects the knowledge in the society. People, especially the young generations listen, watch and learn from the media behaviour. Meanwhile, with the mass migration and return migration in China, migration spreads the knowledge of destination to the native places. Moreover, individuals in villages always exchange knowledge through word-of-mouth communication with their neighbours. The peer effect of neighbourhood knowledge is also modelled. China provides a good case study on how diet knowledge alters through mass media and migration. Mass migration and quick transition of rural China provides a strong incentive to focus on the impact of migration and mass media, the two important channels of information path in rural China, on diet knowledge.

Given the strong effect of diet knowledge on health and human capital in the population, policymakers must understand how mass media and migration changes the diet knowledge. Thus, they will need to regulate the programmes and information delivered in mass media and maximise the social return of migration, e.g., the positive social influences of migration and return migration in villages. Thus, there is strong motivation to know what impacts the diet knowledge of the rural people in a village. Is it media exposure or the migration density in the village or the peer effects? Figure 4.1 shows the causal interrelations that this paper tries to investigate.

Rural residents can communicate, observe and learn from the dietary knowledge of migrants. At the same time, they acquire the information and imitate what they learn from mass media. They also get knowledge from their neighbours. This paper attempts to quantify the externality of migration, the peer effect and the impact of mass media on diet knowledge in the community.

This study also contributes to the emerging literature on the impact of migration,

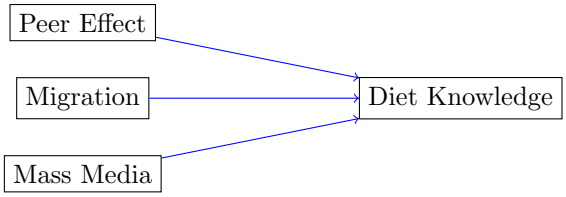


Figure 4.1: Mass media, migration and diet knowledge

particularly on knowledge transmission. It adds to the fast growing literature on the impact of mass media. To the best of the understanding, this is the first empirical work studying the spillover effect of migration on diet knowledge.

The remaining sections of the paper are segregated in the following manner: Section 2 reviews the related studies; Section 3 introduces the data, variable measurements and descriptive statistics; Section 4 presents the research methodology and empirical strategy, presents the main results and interprets the findings; and Section 5 concludes the paper.

4.2 Literature Review

Mass media, including newspapers, the Internet, TV and others, have influenced the human society. Research on the social and economic impacts of mass media has been developing (see a review Prat and Stromberg (2013) on the impact of mass media, regarding the politics and policy, and DellaVigna and Ferrara (2015) for a review on the economic and social impacts of media).

Some studies in the economics literature evaluate the impact of mass media on the social norms, behaviour and belief. Bursztyn and Cantoni (2015) finds that exposure to Western broadcasts affects the consumption behaviour of the residents in former East Germany. Dahl and DellaVigna (2009) indicates that watching violent movies would actually decrease violent crimes. Jensen and Oster (2009) studies the impact of cable TV on women in India and finds that the introduction of cable television decreases domestic

violence, preference for sons and increases women's autonomy. Some studies also focus on the political impact of mass media. Gentzkow (2006) examines the impact of TV on voting turnover and finds significant decreasing results. DellaVigna and Kaplan (2007) discovers that Fox News, a TV news channel in the US influences a section of its viewers to vote for the Republican party. Certain studies also concentrate on the effect of mass media on the studying outcome of children. For instance, Gentzkow and Shapiro (2008) indicates that watching TV would actually increase the school children's test scores.

Besides literature in economics, there is a plenty of evidence regarding the impact of mass media on health and diet in medical research (Wakefield et al. (2010), Roberto et al. (2015) for reviews). In relation to this paper, there are many medical studies on the impact of mass media on health knowledge. Goldberg (1992) finds that audiences are sometimes confused and even reject health recommendations from the media because of the complexity of the information, different perspectives of proponents and certain constraints of media. Chew et al. (2002) finds that TV programmes promoting health can increase health knowledge, enhance health beliefs and contribute to health behaviours. A recent study, H.Nagler (2014), addresses how exposure to conflicting information leads to less belief in health recommendations and less engagement in healthy behaviour.

However, there are no empirical findings regarding the impact of mass media on diet knowledge in the economics literature. This paper investigates this impact, exploiting the variation in the individual mass media exposure and diet knowledge.

Besides contributing to the growing literature on the economics of media, this paper focuses on migration as the other channel with regards to the spread of healthy diet knowledge in the village. Migration influences the migrants and sending areas in many aspects (see Wahba (2014), Dustmann and Görlach (2016) for reviews). For instance, Batista and Vicente (2011) studies the impact of migration on political ideas; Wahba and Zenou (2012) studies the impact of migration on entrepreneurship; and Wahba (2015) studies the impact of migration on labour market outcomes. More specifically, migration can change the behaviour and beliefs of migrants. Some studies in the field of public

health and medicine, such as Chen et al. (2009b), Yang et al. (2007), have examined the impact of migration on health behaviours.

The literature on the impact of migration on social norms, behaviour and belief changes in economics is limited but emerging. Tuccio and Wahba (2015) studies how international migration may transfer gender norms in Jordan. Bertoli and Marchetta (2015), meanwhile, examines the impact of return migration on fertility choice of Egyptian women.

By contributing to this emerging field, this paper studies the impact of migration on knowledge. It especially looks at how migration leads to the diet knowledge spillover in the villages and spreads the knowledge of rural individuals. From existing migration studies, it is seen that migration impacts the migrants themselves. However, this paper tries to understand whether migration has some spillover effect.

In the economics literature, it is a norm to study the diffusion and spillover effect in different areas, focusing especially on technology, innovation and growth (see Jaffe (1986), Barro and Sala-i Martin (1997), and a review for international technology diffusion Keller (2004)). Recent studies include Bloom et al. (2013), studying technology spillover using US firm panel data, Aghion and Jaravel (2015), studying the R & D and knowledge spillover in economic growth process. Apart from these, some papers study the spillover effect in development aspects. Spolaore and Wacziarg (2009) analyses the diffusion of development from world technological frontier, and addresses the importance of the genetic distance as a barrier to diffusion. Banerjee et al. (2013) studies how participation in a microfinance loan programme diffuses via social network in Indian villages. They can examine the mechanism of diffusion, taking advantage of the process of information pass in social networks.

With data available, this paper uses the village migration and individual media exposure to find that migration density and mass media exposure impacts the diet knowledge of rural residents. It is recognised and acknowledged that migration may influence the diet knowledge at different levels. At individual level, the migration experience would change the knowledge of the individual; on the household level, the existence

of migrants or return migrants in the family would change the knowledge of the family members. However, this paper focuses on a more aggregate level–the village level migration density to study particularly the spillover effect. This is one of the innovation and contribution parts of this paper based on existing literature. Furthermore, the whole thesis focuses on migration in China with the second and third chapters aiming to study the direct effect of migration, and the chapter four evaluating the spillover effect, which adds to the diversity of analysis in the thesis. The results show that a one standard deviation increase in the migration rate in the village leads to a 0.02 standard deviation increase in the diet knowledge of rural residents. A one standard deviation increase in mass media of TV increase the diet knowledge by around 0.01 standard deviation, while a one standard deviation increase in the mass media of Internet leads to a 0.02 standard deviation increase in the diet knowledge of rural residents.

4.3 Data

The data used in this paper is a panel survey: the China Health and Nutrition Survey (CHNS) conducted by the Carolina Population Center at the University of North Carolina and Chinese Center for Disease Control and Prevention.

This data includes a household survey, individual survey, nutrition & physical examination and community survey for the year 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and 2011. The entire data includes about 4400 households with $26,000$ individuals across 9 provinces in China. This data provides information on mass media and diet knowledge and questions regarding information on mass media are asked. For instance, do you have access to TV or the Internet from home, relatives or friends? Do you watch TV? If yes, how much time do you spend watching TV? From the information gained from the survey, this paper constructs the measure of exposure to mass media and the time use on mass media.

Apart from questions on the mass media, the knowledge of healthy diet is based on the following questions in the questionnaire: What's your opinion on the diet? Is it good

for your health? Individuals make a decision on whether each of the listed items is good for one's health: diet with fresh fruits and vegetables, eating a lot of sugar, eating a variety of foods, diet high in fat, diet with more staple food, diet with a lot of fish, eggs and lean meat, diet with less fatty meat, consuming milk and dairy products, consuming beans and bean products. The questions are listed in the Table 6.14 in Appendix 6.10.

Individuals respond to the above questions regarding diet by choosing from strongly agree, somewhat agree, neutral, somewhat disagree, strongly disagree. Responses to the questions are coded as follows: disagree or strongly disagree with a healthy diet is coded as 0; neutral with a healthy diet is coded as 1; agree or strongly agree with a healthy diet is coded as 2; Disagree or strongly disagree with an unhealthy diet is coded as 2; being neutral with an unhealthy diet is coded 1 and agree or strongly agree with an unhealthy diet is coded as 0. This paper uses the principal component analysis (PCA) method to convert the set of diet knowledge variables into the principal component, a single index which explains most of the variance and variation of the original set of indexes. The component with the largest eigenvector, and which explains the most of the variation of the original data, is taken as the dependent variable in the models. Descriptive statistics for mass media exposure and diet knowledge are presented in the part 4.3.2.

This paper uses the 2004 and following waves, as both migration and diet questions have been asked since 2004. The sample is restricted to adults who are 20 to 55 years old, as people who are too young, or too old may not have a rationalised and updated idea on the diet knowledge. The sample is also restricted to rural residents because the paper focuses on migration as an important channel of information path in villages.

4.3.1 Variables and Measure

The main variables in this paper are mass media exposure, migration and diet knowledge. The measure $Mass\_tv_{i,t}$ is the exposure to TV for individual i at time t. This paper focuses on the spillover effect of the migration in the village. So the migration variables are measured at the village level to study the externality of migration on the

knowledge of rural residents. The village migration measures are constructed as following: $Mig\_num_{v,t}$ is the number of migrants in village v and time t. $Mig\_hh_{v,t}$ is the number of households with current migrants in village v at time t. Thus, the proportion of migrant households in each village $Mig\_den_{v,t} = \frac{Mig\_hh_{v,t}}{m_{v,t}}$ where $m_{v,t}$ is the total number of households in the village. $Mig\_rate_{v,t}$ is the migration rate in village, which equals to $\frac{Mig\_num_{v,t}}{n_{v,t}}$ with $n_{v,t}$ being the population size of v at time t.

Measures of diet knowledge $Diet\_know_{i,t}$ is the diet knowledge of the individual. The diet knowledge of the individual comes from individual opinions on different types of food. The PCA technique is applied in order to compose the variables into a principal component.

The diet knowledge is the individual opinions on types of diet. The values of the principal component range from -6.87 to 1.66. Mass media exposure is an individual's participation in watching TV. It is 0 or 1. Overall, the TV exposure in Chinese villages is high over the years (details in next part). In the study sample of 19,355 observations, 4,401 (22.74%) report that they "do not watch TV". To accommodate the panel fixed effects models, sufficient within individual variation is required. In the study sample, about 44% individuals have seen changes in the TV exposure over the survey years. The migration variables show that the migration rates and the proportion of migrant households are presented for reasons of robustness.

Table 4.1 shows the main variables, their value ranges and economics meaning.

Table 4.1: Measure of key variables

	Variables	Measure	Range
Diet Knowledge	$Diet\_know_{i,t}$	Diet knowledge of individual	[-6.87, 1.66]
Mass Media	$Mass\_tv_{i,t}$	Individual participation in watching TV	{0,1}
Migration	$Mig\_den_{v,t}$	proportion of migrant households in total households	[0,0.4]
	$Mig\_rate_{v,t}$	percentage of current migrants in village population	[0,0.95]

CHNS, Dependent and main explanatory variables and measurement.

4.3.2 Descriptive Statistics

The data used in this paper ranges from 2004 to 2011. As this paper studies the knowledge of individuals, rural residents aged between 20 to 55 are included in the sample. The sample includes 6,044 individuals in 2,751 households. Individuals are observed at least twice during the waves in order to implement panel fixed effect models. The individual-year observation is 19,355 and each individual is observed about 3 times on average. 2,920 (48.31%) of the individuals are observed 4 times, 1,427 (23.61%) are observed three times and 1,697(28.08%) are observed twice. Due to the no-response on the diet questions, 5,017 observations are discarded, and 2,104 observations dropped as they are observed only once.

Table 4.2 shows the statistics of the main variables. The migration rate is about 12%. It increases from 0.11 in 2004 to 0.13 in 2006, then there is a slight decrease in 2009, and then increases again to 0.14 in 2011. The similar trend is observed for the proportion of migrant households in the village when it grows from 30% in 2004 to 36% in 2006, then decreases to 34% in 2009 and increases to 39% in 2011. The migration rates captured in this paper are the overall migration rates for males and females, children, adults and the ageing population.

The average value of diet knowledge is -0.03. The average diet knowledge is -2.36 in 2004 and increases to 0.74 in 2006 and then stays relatively stable in 2009 and 2011. The average TV exposure of rural residents is 0.77. There is a decrease in village TV exposure from the year 2004 to the year 2006, and the TV exposure increases to 0.79 in 2009 and then decrease to 0.70 in 2011. The average TV exposure for rural residents is high over the years. The time allocation for watching TV is $Mass\_tv\_time_{i,t}$, which is the hours spent watching TV in a week. The average time for watching TV is about 3.26 hours. The Internet access of rural residents is about 0.17 on average. The Internet exposure expands rapidly from 0.07 in 2004 to 0.18 in 2006. The time use on the Internet is about 0.2 hours and increases steadily over time. Figures 4.2 to 4.5 show the variation in the mass media exposure. Figure 4.2 shows the variation in watching TV, while figure 4.3 shows the variation in the time spent watching TV. Figure 4.4 is the

variation of Internet access and figure 4.5 is the variation in time use of Internet. Figure 4.6 is the variation in the diet knowledge.

Table 4.2: Statistics of key variables

Variable	Min	Max	Mean	N	2004	2006	2009	2011
$Mig\_rate_{v,t}$	0.00	0.40	0.12	19,355	0.11	0.13	0.12	0.14
$Mig\_den_{v,t}$	0.00	0.95	0.35	19,355	0.30	0.36	0.34	0.39
$Diet\_know_{i,t}$	-6.87	1.66	-0.03	19,355	-2.36	0.74	0.63	0.65
$Mass\_tv_{i,t}$	0.00	1.00	0.77	19,355	0.83	0.77	0.79	0.70
$Mass\_tv\_time_{i,t}$	0.00	41.50	3.26	19,355	3.61	3.19	3.39	2.87
$Mass\_int_{i,t}$	0.00	1.00	0.17	19,355	0.07	0.18	0.19	0.23
$Mass\_int\_time_{i,t}$	0.00	50.05	0.24	19,355	0.11	0.14	0.35	0.40

CHNS, summary statistics of key variables. Rural residents of age 20 to 55.

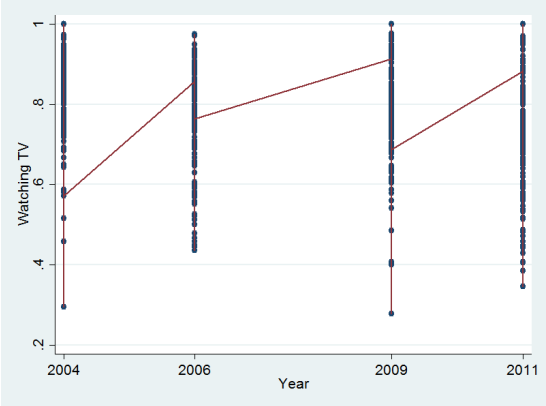


Figure 4.2: Watching TV variation

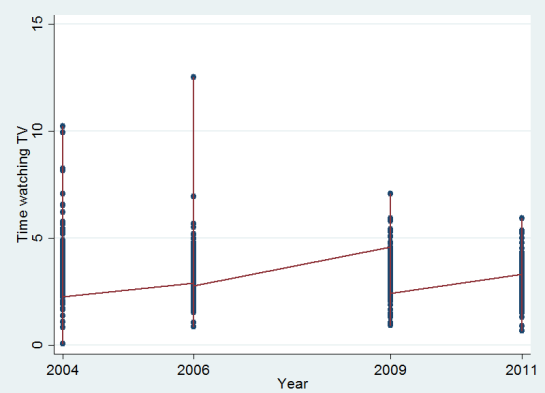


Figure 4.3: Time watching TV variation

The social and economic characteristics at different levels are described in Table 4.3. The sample is restricted to 20 to 55 year old. The mean age of the sample is 39.46 and simply increases over the year. Gender is an indicator of male or female. Minority takes the value 1, otherwise 0. In the fixed effects models, the time-invariant variables are dropped. Education is the level of the highest level of education of the individual. Sick is a dummy variable indicating whether the individual was sick in the past four weeks at the time of the survey. Insurance is a variable that shows whether an individual has

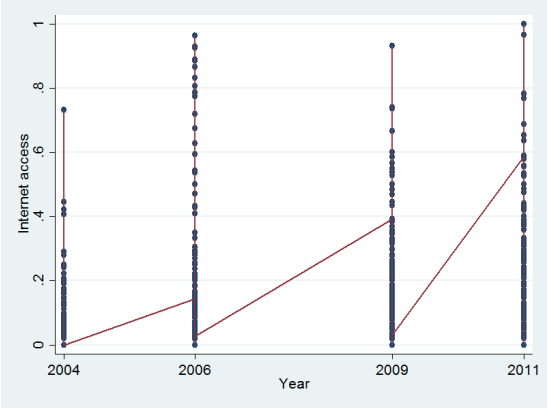


Figure 4.4: Internet access variation

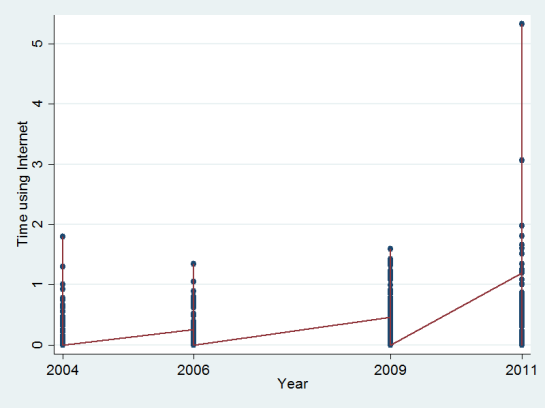


Figure 4.5: Time on Internet variation

any kind of medical insurance. The average medical insurance coverage in rural China grows quickly from about 20% in 2004 to over 90% in 2009.

The household income is the log of household income in 2011 money, and there are some negative values suggesting some rural households may be in debt. The reasons to this are common, e.g. education investments for the children, natural disasters such as flood, drought, or health expenditure. "Pay water" is a dummy for whether the households pays for tap water, which suggests the household has access to clean and processed water automatically at home, instead of water from a well, river or pond. "Toilet facilities" is a dummy for the type of toilet facility available in the household, either it is in-house flush, bathroom or is it a public and open pit. The ownership of a set of household assets such as bikes, ovens, phones is included. The "fuel type" variable is a dummy variable showing the fuel type the household uses for cooking, whether it is coal, electricity, or gas as indicated by 1, and wood, charcoal, or kerosene as indicated by 0. A set of characteristics at village level is included. The population and number of households in each village are the number of surveyed population and households in the sample. These statistics are not real characteristics of the villages but can reflect some facts about the village. This paper takes the advantage of a set of village indexes created by the survey team. These indexes range from urbanisation index to sanitation. For most of the rural development indexes, there is an increasing trend over the years.

Table 4.3: Statistics of control variables

	Variable	Min	Max	Mean	N	2004	2006	2009	2011
Individual	age	20	55	39.46	19,355	38.32	39.39	39.62	40.48
	gender	1	2	1.52	19,355	1.50	1.50	1.51	1.51
	nationality	0	1	0.14	19,355	0.14	0.15	0.14	0.13
	education	0	2	0.88	19,355	0.85	0.89	0.90	0.90
	sick	0	1	0.11	19,355	0.13	0.11	0.12	0.10
	insurance	0	1	0.63	19,355	0.19	0.46	0.90	0.94
Household	household income	-13.00	13.00	9.04	19,355	8.65	8.74	9.28	9.48
	pay water	0	1	0.58	19,355	0.51	0.57	0.59	0.66
	toilet facilities	0	1	0.38	19,355	0.30	0.34	0.42	0.46
	own bike	0	1	0.62	19,355	0.69	0.64	0.57	0.57
	own wash machine	0	1	0.68	19,355	0.56	0.62	0.74	0.82
	own fridge	0	1	0.53	19,355	0.35	0.42	0.62	0.75
	own elc-fan	0	1	0.80	19,355	0.78	0.79	0.81	0.83
	own ovens	0	1	0.14	19,355	0.07	0.11	0.17	0.20
	own elc-pot	0	1	0.72	19,355	0.56	0.68	0.80	0.83
	own phone	0	1	0.59	19,355	0.67	0.64	0.56	0.49
	fuel type	0	1	0.75	19,355	0.67	0.72	0.81	0.81
Village	observed population	3	172	91	19,355	80	92	95	96
	number of households	1	30	21.01	19,355	20.64	20.94	21.29	21.15
	urbanization index	23.39	103.07	58.74	19,355	54.27	56.53	61.13	62.96
	population density	2	9.5	5.63	19,355	5.55	5.62	5.66	5.69
	diversity	2.5	9.5	4.80	19,355	4.23	4.72	5.08	5.15
	economics	0	10	5.70	19,355	5.02	5.72	5.75	6.28
	health	0.8	10	4.91	19,355	4.37	4.30	5.45	5.53
	housing	0.97	10	6.48	19,355	5.63	6.12	6.90	7.26
	market	0	10	3.98	19,355	4.18	3.62	4.16	3.99
	social services	0	10	2.81	19,355	2.16	2.41	3.18	3.49
	transportation	0	10	5.57	19,355	5.54	5.50	5.57	5.66
	education	0.48	7.94	3.00	19,355	2.95	2.97	3.01	3.06
	modern markets	0	10	3.96	19,355	3.95	4.04	3.90	3.94
	sanitation	0	10	5.82	19,355	5.43	5.76	6.02	6.06

CHNS, summary statistics for control variables. Gender: male is 1 and female is 2; Nationality: minority is 1, otherwise 0; Education level: primary or less is 0, middle school is 1, high school and above is 2; Household income: the log of household annual net income in 2011 money; Other variables for household facilities, assets and durable goods are dummies, 1 for yes, 0 for no. Observed population and number of households are for the surveyed individuals and households; Urbanisation to sanitation indexes are investigator's scores for each village, in aspects of education, economic development, transportation, etc. Market is an overall score for markets while modern markets is a valuation for modern function of the markets. Detailed explanation can be found in Appendix 6.9.

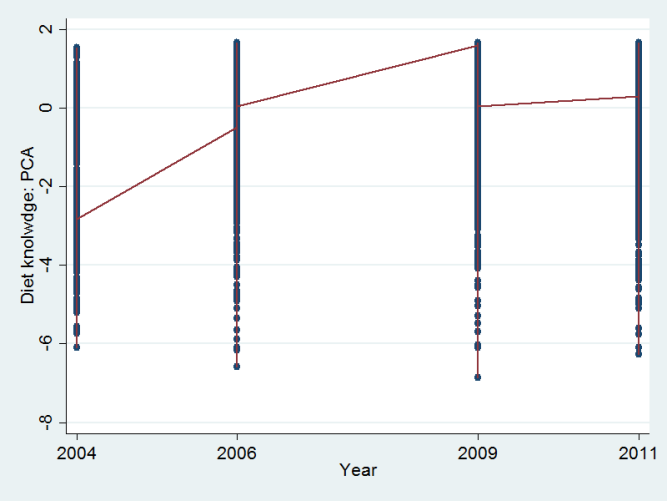


Figure 4.6: Diet knowledge variation over the years

4.4 Empirical Strategy

This paper looks at the impact of mass media and migration on diet knowledge of rural individuals. Figure 4.1 shows the channels of diet knowledge transmission, migration, mass media and peer effect all are important channels for rural residents to receive information. This paper considers all the three channels for studying the diet knowledge transfer in rural China.

4.4.1 Theory Framework

Existing empirical evidence in the economics of international migration (Wahba (2014), Tuccio and Wahba (2015)) indicates that return migrants transfer social norms from migration destination to the native areas. This paper assumes that the return migration would also transfer knowledge from destination to rural areas. Thus, it is expected that the village migration density would have a positive impact on the diet knowledge of rural residents. The healthy diet knowledge from migrants or family members in migrant households would also influence the other residents. Thus, the knowledge also

spreads to the non-migrant rural residents.

According to the findings reviewed in part 4.2, mass media is also an important channel for knowledge transfer in rural China. Exposure to TV or the Internet would bring more information to individuals and change the diet knowledge of rural residents. As the Chinese rural residents are relatively closed and lack advanced diet knowledge, it's expected that mass media would increase the diet knowledge amongst these individuals. Individuals with higher mass media exposure would, hence, have more diet knowledge.

Besides the migration and mass media, there are other ways for the rural residents to acquire information and get knowledge of diet. One way is from their social network. Rural individuals can transfer knowledge via communicating, learning and imitating their family, friends, relatives and other social contacts, especially in rural areas where the transfer of knowledge depends much on the word-of-mouth communication. It's thus expected that, besides the impact of migration and mass media, there is peer effect of knowledge among rural residents. Similar to the evidence in Banerjee et al. (2013), the participation in a microfinance programme in rural India would transmit and diffuse through social networks via word-of-mouth information transfer.

As stated above, the migration density in villages would increase the diet knowledge of rural residents. It is plausible that migrants would transfer knowledge from migration destination to villages, thus villages with a higher migration density would see improved individual diet knowledge. However, individuals are more likely to communicate and exchange knowledge with their social contacts. Thus, it is important to model the peer effect. In this analysis, the social network or information transferring mechanism is not available in the data. It is difficult to explain how the rural residents pass the diet knowledge to each other in the village. This paper captures the peer effect of diet knowledge within the villages by constructing a variable of neighbourhood diet knowledge.

The neighbourhood diet knowledge for individual i is the average diet knowledge of rural residents (excluding i) of the same age group and survey year in the same village with individual i. The rural residents would obtain knowledge from their acquaintances. It's also assumed that individuals would probably communicate and exchange knowl-

edge with individuals of similar age. The neighbourhood diet knowledge is, therefore, expected to have an impact on the diet knowledge of an individual.

4.4.2 Baseline Results

In this study, the two measures of migration in the village are applied for the robustness purpose. The migration measures are the migration rate, which is the percentage of current migrants in village population; and migration household density, which is the proportion of migrant households in total households of the village.

The mass media measures can be the exposure to TV measured intensively or extensively, i.e. the participation in watching TV or the time spent on watching TV. It would be interesting to study how the time spent watching TV impacts the diet knowledge of individuals. This exercise studies the time use data on watching TV and attempts to find evidence that, if more time is spent watching TV, it would improve the diet knowledge of the rural residents. The time spent watching TV is a self-reported number of hours spent on watching TV in a week.

As for the mass media exposure, this paper considers the Internet as well. For the robustness of the results, it is reasonable to consider different forms of mass media. The world is developing in many aspects including the pattern of mass media. As a form of modern mass media, the Internet is quicker, user-interactive and delivers more information. It's necessary to study whether the Internet, as a form of modern mass media, has an impact on the diet knowledge. The Internet exposure of an individual is an individual's access to Internet via various channels like home-owned computers, from friends or relatives or charged Internet service. The other measure of Internet exposure is the time spent on the Internet. The time spent on the Internet is a self-reported number of hours spent on the Internet in a week.

Overall, this paper uses two different measures of village migration to show the robustness of the results, and also uses two different mass media, TV and Internet, in their intensive and extensive margins to provide robust findings. In order to facilitate

interpretation and comparison of the results, this paper reports standardised coefficients for all the regression results.

The linear fixed effects model with $Peer_{i,t}$, the neighbourhood diet knowledge of individual i at time t is implemented. To test the hypothesis, this paper takes advantage of panel property of the data and applies the fixed effects model to control for any time invariant unobserved factors that may impact diet knowledge of the rural residents. The diet knowledge of individual i at time t is modelled as follows:

$$Diet_{i,t} = \theta_t + \beta_1 Mass_{i,t} + \beta_2 Mig_{v,t} + \beta_3 Peer_{i,t} + \gamma X_{i,t} + c_i + u_{i,t} \tag{4.1}$$

The diet knowledge is $Diet_{i,t}$. The mass media exposure measurements $Mass_{i,t}$ can be the participation in watching TV, the time spent on watching TV, the participation in Internet or the time spent on Internet. Migration $Mig_{v,t}$ is $\{Mig\_den_{v,t}, Mig\_rate_{v,t}\}$, proportion of migrant households or percentage of migrants in the village population.

$X_{i,t}$ is the set of controls of individual, household and village level characteristics in Table 4.3, a detailed explanation and describe of the variables can also be found in Appendix 6.9. One particular variable is the education of the rural individuals. Although it is reported in the summary statistics, in the analysis this paper creates three dummy variables, each of them stands for one education category - primary education, middle education and higher education. The aim is to compare the impact of mass media, migration and the education to check if mass media and migration do really affect the diet knowledge. The regression models treat the primary education as the reference group and report the standardised coefficients of middle and higher level education.

c_i is the individual fixed effects. In this model, arbitrary correlation of c_i with $Mass_{i,t}$, $Mig_{v,t}$, $Peer_{i,t}$ and $X_{i,t}$ is allowed to deal with any time invariant unobserved effects. A time dummy variable θ_t is included. $\beta_1, \beta_2, \beta_3$ and γ are the parameters respectively, $u_{i,t}$ is the error term.

The reverse impact issue is not relevant as individual diet knowledge does not have a direct effect on the mass media exposure or the village level migration. The analysis of diet knowledge starts from the linear panel fixed effects model, introducing the peer

effect in the models, and the two different mass media, TV and Internet are introduced simultaneously for comparison reasons.

The baseline estimation of the impact of migration and mass media is implemented with a linear panel fixed effects model with the regression equation 4.1. The dependent variable is the individual diet knowledge. The individual mass media, village migration measurements and neighbourhood diet knowledge are included. The individual, household and village level control variables, together with individual fixed effects and time fixed effects are modelled. The estimation results for TV are reported in Table 4.4 and the results for Internet are presented in Table 4.5.

The baseline models in Table 4.4 and Table 4.5 both consist of 4 models. In Table 4.4, model 1 and model 2 show the impact of village migration and mass media on the individual diet knowledge when the TV measure is participating in watching TV, while model 3 and model 4 are models when TV measure is time spent on watching TV. Similarly, in Table 4.5, model 1 and model 2 show the impact of village migration and mass media on the individual diet knowledge when the Internet measure is participating in using the Internet, while model 3 and model 4 are models when Internet measure is time spent on the Internet.

In all the eight baseline models, the significant F statistics and adjusted R-squares show that the overall model fits well and has good explanatory power. The robust standard errors are clustered at the village level. The knowledge peer effect is modelled. The education dummy variables for middle or high level are presented in the models, treating the primary education as the reference group. The standardised coefficients of exposure to and time spent on TV and the Internet are presented in Tables 4.4 and 4.5 respectively, in order to make it easy to evaluate whether the magnitude of the impact differs with the media.

Table 4.4: Diet knowledge: baseline fixed effects models for TV

	Model 1	Model 2	Model 3	Model 4
Village migration rate	0.015**		0.014**	
	(0.007)		(0.007)	
Proportion migrant household		0.015**		0.014**
		(0.007)		(0.007)
TV exposure	0.012***	0.012***		
	(0.004)	(0.005)		
TV time			0.004**	0.004**
			(0.002)	(0.002)
Knowledge peer effect	0.604***	0.604***	0.603***	0.603***
	(0.026)	(0.026)	(0.026)	(0.026)
Education-Middle	0.010	0.010	0.011	0.011
	(0.011)	(0.011)	(0.011)	(0.011)
Education-High	0.003	0.003	0.005	0.005
	(0.013)	(0.013)	(0.013)	(0.013)
Individual controls	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes
Village controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
F	39.29***	39.78***	38.62***	38.95***
Adj R2	0.74	0.74	0.74	0.74
N	19,355	19,355	19,355	19,355

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Cluster robust standard errors in parentheses.

Table 4.5: Diet knowledge: baseline fixed effects models for Internet

	Model 1	Model 2	Model 3	Model 4
Village migration rate	0.014**		0.014**	
	(0.007)		(0.007)	
Proportion migrant household		0.014**		0.015**
		(0.007)		(0.007)
Internet exposure	0.015***	0.015***		
	(0.005)	(0.005)		
Internet time			0.015***	0.015***
			(0.005)	(0.005)
Knowledge peer effect	0.595***	0.595***	0.596***	0.596***
	(0.026)	(0.026)	(0.026)	(0.026)
Education-Middle	0.011	0.011	0.011	0.011
	(0.011)	(0.011)	(0.011)	(0.011)
Education-High	0.002	0.002	0.001	0.001
	(0.013)	(0.013)	(0.014)	(0.014)
Individual controls	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes
Village controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
F	37.22***	36.89***	37.85***	37.58***
Adj R2	0.73	0.73	0.73	0.73
N	19,355	19,355	19,355	19,355

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Cluster robust standard errors in parentheses.

As expected, the coefficients of village migration (migration rate and proportion of migration household in the village) are significant and positive. One standard deviation increase in the village migration rate would improve the diet knowledge of individual "i" in the village by about 0.014 standard deviation (Models 1 and 3 in Tables 4.4 and 4.5 for the village migration rates, Models 2 and 4 in Tables 4.4 and 4.5 for proportion

of migrant households, and the coefficients of migration measurements are consistent in Tables 4.4 and 4.5). This result suggests that the migration increases the knowledge of rural residents.

Two different measures of migration in the village are separately controlled to show the strength of the findings. The impact of migration on the diet knowledge is robust to different measures of migration. The models show a positive and significant impact of migration on the individual diet knowledge, irrespective of what migration measurement is included. As internal migration is rural to urban migration, the migration destinations are more modern and developed as compared to the villages. Migrants bring diet knowledge back to native villages and then the knowledge spillover to rural residents occurs. The findings in this paper are consistent with the empirical evidence that migrants transfer the social norms and beliefs in the literature of international migration.

The models with TV exposure (Models 1 and 2 in Table 4.4) are overall significant. A one standard deviation increase in the TV exposure would increase the diet knowledge of the individual by 0.012 standard deviations. The effect of time use in watching TV is positive and significant (Models 3 and 4 in Table 4.4). One standard deviation increase in time spent on watching TV leads to 0.004 standard deviation increase in the diet knowledge. This exercise also provides evidence that the TV exposure in either intensive or extensive margin, would improve the diet knowledge when village migration is in different measurements.

Participation in watching TV would expose the rural individuals to more information and knowledge on diet. More time spent watching TV means the individuals would receive more information. The impact of mass media on diet knowledge is unclear in theory and previous empirical works. However, this paper shows evidence that both the incidence of watching TV and the time used for watching TV would increase the diet knowledge of the individuals.

The regressions of the linear fixed effects models for internet access and time use on the internet are presented in Table 4.5. TV is replaced with internet measures and the same regression models for equation 4.1 are implemented. The results show that the

coefficients of internet access and internet exposure time are all positive and significant. A one standard deviation increase in the internet access would improve the diet knowledge by around 0.015 standard deviation and one standard deviation increase in the time spent on the internet increases the diet knowledge by 0.015 standard deviation.

The results show that measurements of internet use have a significant impact on the diet knowledge. The Internet is a growing channel (Figure 1.6), from which rural residents can gain information about the outside world. Not only access to the internet but also the longer time expose to the internet would have a positive impact on the diet knowledge of rural residents. The internet as a form of mass media transfers the diet knowledge to rural residents. This finding is consistent with the results when the mass media exposure is measured as TV exposure, while the impact of Internet is larger than the impact of TV, in both intensive (0.015 vs 0.004) and extensive (0.015 vs 0.012) aspects.

The peer effect, which is the spillover of neighbourhood knowledge, is positive and significant in all the eight models in Tables 4.4 and 4.5. A one standard deviation increase of neighbourhood diet knowledge leads to 0.6 standard deviation increase in the individual diet knowledge. This effect is relatively large. The knowledge exchange between rural residents of similar age would increase the diet knowledge of individuals. The impact of neighbours is much larger than the impact of mass media and migration, even though the mass media and migration both have positive and significant impacts. The peer effect is large in magnitude. The coefficients of peer effect are greater and consistent when the mass media is measured as TV (Table 4.4) or the Internet (Table 4.5) and when the village migration is measured as migration rates or proportion of migrant household.

Education is an important factor for individual knowledge. The tables show the coefficients of middle-education (a dummy shows the category of education of the individual at middle school level) and high-education (a dummy shows the category of education at high school level or above). The regressions treat the primary education as the reference group and the results show that education does not impact the diet knowledge.

This result is not surprising, as knowledge does not mean diet or healthy knowledge. Literacy and knowledge can be specific to many other areas - math, science, environment and so on. A recent study in health and medicine Johnston et al. (2015) shows that increased schooling does not significantly affect health knowledge and future research is recommended to determine whether there is a causal link between education and health knowledge.

The largest human mobility in history, internal migration from rural to urban China, results in many changes to the rural areas of China. One of the influences is the transfer of knowledge. Migration spreads the knowledge of diet in the village and increases the individual diet knowledge. Migrants are linkages between the rural villages and urban areas, where knowledge of diet is more available and accessible. The results suggest that the diet knowledge in the urban areas can be diffused into the home villages via migration.

Beyond migration, the rural residents also obtain information outside of the village via other channels. They can benefit from the expansion of mass media to have more information about the outside world. Individuals can obtain more knowledge on many aspects, including diet. The finding in the baseline model provides evidence that more exposure to mass media is associated with more diet knowledge. The baseline models also show that there is a significant and large peer effect of knowledge, while the education does not have a significant effect on the diet knowledge.

4.4.3 Endogeneity

It is always challenging to identify the causal effect in regression models. It's vital to understand that the diet knowledge of individuals does not influence the migration decision of the individuals. Furthermore, the diet knowledge of the individuals does not directly impact the village migration. Moreover, the migration density of the village in the model is not endogenous.

Also, TV programmes delivered can be numerous. H.Nagler (2014) finds that infor-

mation on the TV programmes can lead audiences to become confused about the healthy diet. Some programmes offer specialised recommendations on healthy diet, which are expected to increase the diet knowledge of the audience. Moreover, some programmes, such as advertising of fast food and high sugar soft drinks would have a negative effect on the diet knowledge, especially in the case of the young generation. This paper has provided evidence that the exposure to TV and the time spent on watching TV has a positive and significant effect on the diet knowledge of the individuals. However, TV exposure is endogenous in the analysis and the estimation of the linear panel fixed effects models can be biased.

The endogeneity of the individual TV exposure may come from many sources. People of some typical characteristics may choose to watch more TV programmes rather than participating in other activities. These personal characteristics are difficult to measure and thus omitted in the model. The individual fixed effects model would control for some individual level characteristics that may lead to the issue of endogeneity.

There may exist some time-changing unobserved factors that affect the TV exposure and have a direct impact on the diet knowledge of the individuals. For instance, individuals who care more about diet would choose to watch programmes about professional comments on dietary information, and the concerns about diet affect many other issues, such as the physical or physiological situation of the individuals. An individual's attitude towards the information from mass media would also change over time. These issues make the baseline models quite suspicious.

This paper proposes an instrumental variable that is closely related to the individuals TV exposure but does not directly impact the diet knowledge of the person. The IV applied for individual i at time t is village TV exposure, which is constructed as the average TV exposure of the people of the same gender and same age group in the same village at a particular time t.

The village TV exposure is closely correlated with the TV exposure of an individual, especially when the village TV exposure is constructed from individuals of the same age and gender group. An individual would be more likely to watch TV if the rural

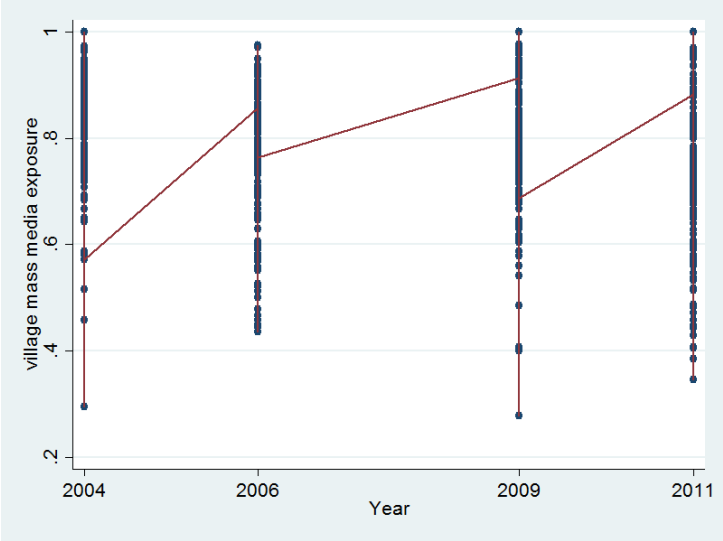


Figure 4.7: Mass media exposure in villages over the years

residents in the same village engage in the same activity. This is because people within the network tend to participate in same activities, such as watching TV and discussing the same TV programmes. Thus, the aggregate TV exposure can predict the individual behaviour of watching TV and it is exploited to instrument the TV exposure of the individuals. There are abundant variations in the instrumental variable, both among individuals and across time. Figure 4.7 shows the variation of the IV. A scatter point depicts the mean TV exposure in a village. The IV has more variation as individuals of different age and gender would each have a different value for the IV. Every point in the graph represents the IV for a group of individuals observed in the same gender and age group in the same village and year.

The village TV exposure does not impact the knowledge of the individuals directly. In fact, it offers a more aggregated impact on the knowledge of the individual via the individual mass media exposure. Thus, the village mass media exposure is excluded from the model. It might be challenged that the village TV exposure would impact the individual diet knowledge. This is reasonable, as there is an externality of mass media. The village TV exposure influences the individual diet knowledge via two channels, diet knowledge in the village and the individual mass media exposure.

The fact is the village TV exposure may influence the diet knowledge in the village and then the diet knowledge in the village impacts the diet knowledge of individuals. This channel of the knowledge transfer is actually directly captured by the knowledge peer effect. Thus, the indirect channel that the village mass media may influence the individual diet knowledge is measured and controlled in the model. In other words, village TV exposure has two channels to impact the individual diet knowledge; one is through individual mass media exposure, the second channel is through the village mass media exposure that impacts the village diet knowledge, such that the village diet knowledge spill over to the individuals in the village. The second channel is captured by the peer effect of diet knowledge in the model. Thus, in the model setting, village mass media only impacts the individual diet knowledge via the individual mass media exposure.

The instrumental variable strategy is combined with individual fixed effects and time fixed effects. The models are implemented by a two-stage process and all standard errors are clustered at the village level to control for any correlation within the village. Regression results are presented in Table 4.6. Similar to the individual TV measures, the exposure to the internet and time use on the internet are also endogenous. The results of the linear fixed effects models are biased. This paper proposes an instrumental variable to control for the endogeneity of the internet measures, following the logic in the IV construction for TV exposure. The IV for the internet is the village internet exposure, which is constructed for individual i as the average internet exposure of rural residents of the same age and gender group with individual i in the same year and village. The FEIV regression results for the internet measures are presented in Table 4.7.

Table 4.6: Diet knowledge: FEIV models for TV

	Model 1	Model 2	Model 3	Model 4
First Stage				
Village TV exposure	2.357***	2.357***	3.987***	3.983***
	(0.006)	(0.006)	(0.066)	(0.066)
First stage F-statistics	17,000	17,000	3639.38	3643.74
Prob > F	0.000***	0.000***	0.000***	0.000***
Second Stage				
TV exposure	0.013***	0.013***		
	(0.005)	(0.005)		
TV time			0.008***	0.008***
			(0.003)	(0.003)
Migration rate in village	0.015**		0.015**	
	(0.007)		(0.007)	
Proportion of migrant household in village		0.015**		0.015**
		(0.007)		(0.007)
Knowledge peer effect	0.604***	0.604***	0.602***	0.602***
	(0.026)	(0.026)	(0.026)	(0.026)
Education-Middle	0.010	0.010	0.010	0.010
	(0.011)	(0.011)	(0.011)	(0.011)
Education-High	0.003	0.003	0.003	0.003
	(0.013)	(0.013)	(0.013)	(0.013)
Overidentification test	Exactly	Exactly	Exactly	Exactly
	Identified	Identified	Identified	Identified
Underidentification test	122.59***	122.48***	119.71***	119.82***
Weak identification test	17,000***	17,000***	3639.38***	3643.74***
Weak instrument robust inference	7.45***	7.26***	7.45***	7.26***
Durbin-Wu-Hausman test	8.47***	8.35***	4.38**	4.11**
Individual Controls	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
N	19,355	19,355	19,355	19,355

CHNS 2004-2011. $^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Cluster robust standard errors in parentheses.

Table 4.7: Diet knowledge: FEIV models for Internet

	Model 1	Model 2	Model 3	Model 4
First Stage				
Village Internet exposure	2.902***	2.902***	1.517***	1.517***
	(0.010)	(0.010)	(0.102)	(0.102)
First stage F-statistics	86,774	86,851	220.27	221.10
Prob > F	0.000***	0.000***	0.000***	0.000***
Second Stage				
Internet exposure	0.016***	0.015***		
	(0.006)	(0.006)		
Internet time			0.032***	0.032***
			(0.011)	(0.011)
Migration rate in village	0.014**		0.015**	
	(0.007)		(0.007)	
Proportion of migrant household in village		0.014**		0.016**
		(0.007)		(0.007)
Knowledge peer effect	0.595***	0.595***	0.595***	0.595***
	(0.026)	(0.026)	(0.026)	(0.026)
Education-Middle	0.011	0.011	0.010	0.010
	(0.011)	(0.011)	(0.011)	(0.011)
Education-High	0.002	0.002	-0.003	-0.003
	(0.013)	(0.013)	(0.014)	(0.014)
Overidentification test	Exactly Identified	Exactly Identified	Exactly Identified	Exactly Identified
Underidentification test	86.83***	86.80***	70.05***	70.07***
Weak identification test	86,774***	86,851***	220.27***	221.10***
Weak instrument robust inference	7.24***	7.21***	8.09***	8.04***
Durbin-Wu-Hausman test	7.81***	7.64***	4.25**	4.20**
Individual Controls	Yes	Yes	Yes	Yes
Household Controls	Yes	Yes	Yes	Yes
Village Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
N	19,355	19,355	19,355	19,355

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. Cluster robust standard errors in parentheses.

In Tables 4.6 and 4.7, the dependent variables are the diet knowledge of rural res-

idents for the models. Models 1 and 2 are the models in which the individual mass media is measured as mass media exposure (TV in Table 4.6, the Internet in Table 4.7), and models 3 and 4 uses the time allotted to the mass media. Models 1 and 3 use the migration rate as village migration measure, while models 2 and 4 use the proportion of migrant households. Each model includes time fixed effects and individual fixed effects. The control variables are described in Table 4.3 and more explanation is provided in Appendix 6.9.

In the first stage regressions, when the individual mass media exposure is the dependent variable, coefficients on the instrument variable (village TV exposure in Table 4.6, village Internet exposure in Table 4.7) are positive and significant. The first stage F values of the models are large and significant. The first stage statistics show a strong correlation between village mass media exposure and individual mass media exposure. The models are exactly identified. The instrumental variable works well according to the set of statistics. As the instrumental variable methods are applied in these regressions, the test of endogeneity of the endogenous variable individual mass media has been implemented and reported in the Tables 4.6 and 4.7. The test implemented is the "Durbin-Wu-Hausman test of endogeneity" (Durbin (1954), Hausman (1978),Wu (1974)) and reports the Durbin-Wu-Hausman Chi Square test. The results show that the individual mass media should be treated as an endogenous variable.

The second stage coefficients of mass media exposure and migration measures in Table 4.6 are positive and significant. In model 1, the coefficient of migration rate is 0.015 and it is 0.013 for individual TV exposure. In model 2, the coefficient for the proportion of migrant households is 0.015 and significant. For TV exposure, the coefficient remains 0.013. In model 3, the coefficient for village migration rate remains 0.015, and for time spent on TV, it is 0.008. In model 4, the coefficient for the TV exposure time remains the same as in model 3 and the impact of migration density is consistent with model 2. In the second stage regressions in Table 4.7, the Internet access and time exposure on the internet both have a significant and positive effect on the individual diet knowledge with positive standardised coefficients 0.015 and 0.032, respectively. The coefficients of the migration measures are consistent with those in

Table 4.6.

The peer effect of diet knowledge is 0.6, which is consistent with the peer effect captured in the linear fixed effects models. The peer effect captures the influence of the neighbourhood in the villages on the knowledge. This is one of the channels through which the village mass media influences the individual diet knowledge indirectly. The impact of education is not significant and is similar to what was presented in the baseline model results of Tables 4.4 and 4.5. This exercise deals with the endogeneity of individual mass media exposure. The results of the FEIV regressions provides evidence that the individual mass media exposure of TV or Internet and the village migration density has a significant and positive effect on the diet knowledge of rural residents.

It is noticed that the coefficients of individual mass media increased when compared to the linear fixed effects models in Tables 4.4 and 4.5. For the individual TV exposure, the linear fixed effects models show that a one standard deviation increase in TV exposure would lead to about 0.012 standard deviation increase in diet knowledge. However, the instrumental variable fixed effects model suggests the increase should be around 0.013 standard deviation. As for the individual TV exposure time, the linear fixed effects models show that one standard deviation increase in watching TV would increase the diet knowledge of the people by 0.004 standard deviation, while in the FEIV regressions the number is 0.008. These suggest that the linear panel fixed effects models underestimate the impact of the individual mass media exposure on the diet knowledge. The FEIV regression corrects the bias in the linear fixed effects models. Similar with the FEIV regressions for the TV in Table 4.6, the coefficients for mass media measures in FEIV regressions in Table 4.7 are larger than the ones in the linear fixed effects models. For the internet access, the coefficient increases are around 0.015 to 0.016. For the time use on the internet, the numbers increase from 0.015 to 0.032. The instrumental variable corrects the biases in the linear models.

Overall, the impact of individual mass media exposure, village migration and peer information on diet knowledge are significant. The results are consistent with the baseline models. However, the FEIV results show a larger impact of mass media exposure

after correcting the bias of linear fixed effects models. The FEIV results for mass media exposure are slightly larger than the ones in baseline results and the FEIV results for the time use in mass media are about 2 times the ones in baseline results.

4.5 Conclusion

Mass media and migration are reshaping the rural society in many ways in developing countries. This paper studies how diet knowledge spreads across the village through migration, mass media and peer effect.

To answer the research question, this paper uses panel data on the diet knowledge of 19,355 individuals in 146 villages distributed across 9 provinces in China and exploits the time variation of the migration density in the villages and individual mass media exposure. The fixed effects models find that migration and mass media both have a strong positive impact on the diet knowledge of rural residents.

To deal with the endogeneity of mass media, the village mass media exposure is exploited to an instrument for individual mass media exposure. When the peer effect of neighbourhood knowledge is controlled in the model, the IV can only impact the individual diet knowledge via the individual mass media exposure. The two-stage regressions with fixed effects show that the individual mass media has a positive impact on the diet knowledge, while village migration density has a strong spillover effect and transforms the diet knowledge of rural residents.

To make the results more robust, this paper uses the Internet exposure measurements in place of the TV exposure. The Internet exposure and time use of the Internet are examined. The fixed effects results are similar to that of TV exposure, and prove that the Internet exposure has a strong effect on the diet knowledge. The endogeneity of the Internet is also dealt with by the instrumental variable approach. The results show a positive and significant impact of Internet, while the linear fixed effects models underestimate the impact of the Internet. This is consistent with the findings in the models

when TV is the mass media measure.

The intensive margin of mass media exposure is also exploited. The increase in time spent watching TV or on the Internet would lead to an increase in diet knowledge. Besides the impact of the migration density and mass media exposure, this paper indicates the existence of a strong peer effect of neighbour knowledge.

This paper contributes in several ways. It is the first study on the spillover effect of migration on the knowledge in rural areas. This paper quantifies the impact of mass media and migration on diet knowledge, adding to the literature on the transferring of social norms via migration and the literature on the influence of social media.

It's important to address that the study outcome of this paper is the health diet knowledge, not actually the health behaviour or health outcomes. It is true that diet knowledge differs from the behaviour and health status of the individuals. However, there is abundant literature, especially in the field of public health, proving that the health diet knowledge is strongly and positively correlated with the health behaviour and outcomes, for instance the studies reviewed in this paper (Strauss and Thomas (1998), Blaylock et al. (1999), Chew et al. (2002)), and more recent studies (Spronk et al. (2014): Better nutrition knowledge is also associated with better dietary intake; Shojaeizadeh and Gashtaee (2015): There is a positive correlation between knowledge, attitude and healthy behaviour). It is clear that diet knowledge is related to health outcomes. This paper plans to focus on the transfer of knowledge, which is in-line with the emerging research in the transfer of norms on international migration. Furthermore, to show more correlation between the health diet knowledge and the health behaviour using the available information, Appendix 6.11 presents correlation matrices of the diet knowledge and the smoking and drinking behaviour (which are considered an unhealthy behaviour). The negative correlation (Table 6.15) shows that diet knowledge is associated with diet behaviour, and more specifically negatively correlated with unhealthy behaviour. The regression table (Table 6.16) in Appendix 6.11 presents the results of regressions of the smoking and drinking outcomes on the diet knowledge. The negative and significant coefficients also provide evidence that diet knowledge is negatively associated with the

unhealthy behaviour.

Considering the strong effect, it is important for policymakers to address the social media exposure in the developing areas. The information and knowledge acquired from the media are important for the rural residents. In this paper, strong evidence proves that exposure to mass media can improve the diet knowledge. To have access to more TV and the Internet would predict a significant increase in the diet knowledge of the rural residents. These findings would also offer some policy suggestions on public health, This is because knowledge of diet is the starting point of diet choice and behaviour, which are important determinants of health for the rural residents. Head et al. (2015) finds that mass media reduces the child mortality by $10 - 20\%$ and suggests that the importance of mass media should be addressed more while promoting public health.

In addition, the migration and return migration are important knowledge channels for rural development. Taking advantage of the positive externality of migration would enhance the welfare of migrants and non-migrant rural residents. The evidence in this paper indicates a strong impact of migration on the healthy diet knowledge in the rural areas. The economic impact of migration is verified in the economic literature. A spillover effect of migration on the social norms and knowledge rural area, such as health beliefs, female status, environmentally friendly ideas, etc., can be a topic of future studies. This paper contributes to this emerging field and provides evidence of a positive spillover effect of migration on the diet knowledge. Policy makers would encourage this knowledge-acquisition from migrants and return migrants in order to utilise the positive externality for increasing the social returns of migration.

5

Conclusion

This thesis examines the impact of migration in villages during the process of dramatic urbanisation and quick transition period in China. The rural to urban migration is the largest population movement in the human history. The migration and return of numerous rural labourers between the rural areas and the urban areas, especially the coastal cities, have brought about varied changes to both the rural and urban regions. Due to the household registration system, amongst many other reasons, migrants experience constraints when accessing the health, pension, education and other public services, which are only available to urban residents with urban registration. Children and spouses of the migrants are always left behind in the villages and the migrants themselves have to return there. Because of the economic development needs and population structure change (Figure 1.5), there have been many reforms in the household registration policy [1]. Recently, the local and central government of China has been encouraging the rural

[1] https://www.gov.uk/government/publications/hukou-reform-central-government-sets-out-a-vision-august-2014/hukou-reform-central-government-sets-out-a-vision-august-2014

residents to settle in small and medium cities near their villages in order to enhance urbanisation and, boost consumption. However, the reform plans of the "Hukou system" in major migration destinations, especially in coastal cities, are primarily still on paper and only a few have been put into practice. There is still a limitation for the rural residents in the urban areas with respect to health services, education and other services. Many children and spouses are left behind in villages, and migrants still, are compelled to return.

With this background, the thesis focuses on the impact of migration on the rural residents and rural areas. Migration influences the migrants, return migrants and the left behind individuals in many aspects. This thesis comprises of three papers that analyse the impact of migration on the time utilisation of the left behind, the effect of return migration on the smoking and drinking behaviour, and the spillover effect of migration on the diet knowledge of rural residents. This thesis uses the panel data of the China Health and Nutrition Survey for the empirical analysis.

Migration influences the left behind family members through income and substitution effects. The income effect that migration generates is higher incomes. This leads to a decrease in the time use and labour supply of the left behind. However, the substitution effect works contrarily. The left behind may have to increase the time use in order to compensate for the lost labour within the household because of migration. The first paper evaluates this net effect of migration on the time use of the left behind children and spouse. A fixed effects instrumental variable strategy is applied to deal with the time variant and invariant sources of endogeneity of migration. The participation and time use on three types of activity-wage labour, housework and agriculture are studied, for both the left behind spouse and children of male migrants. The results suggest that the income and substitution effects cancel out to make the net effect of migration on the left behind statistically insignificant. One exception from the primary findings is that the father's migration increases the children's time use in wage labour, but this impact is only significant for boys. Boys may have to work in paid jobs when the family faces short-term financial difficulties because of the migration costs or the delay in receiving remittances from the migrants. However, this impact vanishes when a robust standard

error is clustered at village level and the results become consistent with the main findings, which state that the impact of migration on the time use of the left behind is not significant. The empirical result suggests that the two effects trade off with each other.

The Chinese government wants to boost consumption to stimulate the economy. The vast rural markets and mass migrants and return migrants with more capital are seen as a new engine for growth. The second paper of the thesis is the first empirical work to study the consumption behaviour of return migrants. The study focuses on the consumption of tobacco and alcohol, which is common in China and is related to many other personal aspects like health, savings etc. The study applies two empirical strategies to quantify the effect of migration on alcohol and tobacco consumption. The first strategy is the instrumental variable approach dealing with the selection of migration, while the second is the difference in difference (DiD) setting for comparing the before and after migration smoking and drinking behaviour. The study provides evidence that male returnees smoke and drink more than male non-migrants, while for rural female returnees, the results show that their smoking and drinking behaviour is not statistically dissimilar than the female non-migrants. Basic DiD models and the DiD with IV model, all provide strong evidence that there are significant before and after migration changes in smoking and drinking behaviour for males returnees. The results are consistent when the income measure is controlled. This suggests that the smoking and drinking behaviour alteration of male returnees is not simply a result of the income effect. The results provide suggestive evidence that male migrants exposed to the urban norms bring back their smoking and drinking behaviour upon their return to the villages.

The results indicate that the migrants have a significant impact on the consumption of tobacco and alcohol. The third paper aims to determine whether migration has a spillover effect in the concerned villages. The studies illustrate that the village migration density impacts the diet knowledge of rural residents. In fact, it is reshaping the rural society in numerous ways in developing countries. It is an important channel through which rural residents can be educated. The other important channel of information distribution is the mass media. The study models the impact of individual mass media exposure and village migration density on the diet knowledge. To deal with the

endogeneity of individual mass media exposure, the village level mass media exposure is exploited as an instrumental variable. The instrumental variable and fixed effects models indicate that the mass media has a strong impact on the diet knowledge of rural residents. The results also depict a strong spillover effect of migration on the diet knowledge and a significant peer effect. It has also been noted that the linear fixed effects models would underestimate the impact of mass media on the diet knowledge.

The thesis innovates in numerous aspects. In the technique aspect, the thesis uses the panel data exploiting the time variation of the individual observations. Novel instrumental variables are proposed to deal with the self-selection of migration and the endogeneity issues in the model. There has been limited research on migration, primarily with the focus on rural residents and villages in China. This thesis is different from traditional research on the subject in China, which always focuses on the migrants in urban areas. This thesis is also innovative in the way it studies the impact of migration and return migration, which is conventionally under-studied and is beyond the traditional domain of migration studies in the economics literature.

This thesis contributes to the existing research on the impact of migration and return migration, especially the emerging empirical works on how migration transmits social norms and belief. It provides evidence that internal rural to urban migration has a considerable impact on the rural areas and the rural residents, especially on the consumption behaviour and the diet knowledge of rural residents. The first paper in the thesis quantifies the net effect of migration on the left behind and contributes to the field by settling the empirically unclear income effect and substitution effect in existing literature. The second paper is the first study on the consumption behaviour of male returnees and the findings indicate that migration changes the smoking and drinking behaviour. This thesis provides evidence that migration transforms the consumption behaviour of rural males and contributes to the research on the impact of return migration. The third paper is the first study on the spillover effect of migration on knowledge of rural individuals. This paper also quantifies the impact of mass media on the diffusion of diet knowledge, adding to the literature on the transferring of social norms via migration and on the influence of social media.

This thesis also has certain limitations. Firstly, the study outcomes reflect just certain aspects of the rural residents - their time use, smoking and drinking consumption or diet knowledge. Hence, it cannot provide comprehensive inference about the overall impact of migration on the rural area and residents. The influence on other aspects, such as the health, nutrition and education, can be inferred from the results, but the thesis offers no direct evidence.

Nonetheless, the thesis still has positive policy implications. For instance, public funds can set up a micro credit to finance a small loan in order to solve the short-term financial difficulties after migration. Regarding the smoking and drinking consumption, the policy makers can try to inform the public about the health hazards and financial risks of smoking and drinking, especially for migrants who have limited knowledge about health and long-term financial planning. The positive impact of migration on consumption would suggest that the return migrants would want to consume more. The Chinese government's current plan to encourage migrants to buy flats in small and medium cities may also work. However, the authorities may need to encourage the migrants to use their accumulated capital rationally, such as investing in education or skill training, to improve human capital, to finance agriculture or business activities and so on, instead of consuming for the current period but losing the long-term utility. Further, the Chinese government's plan to drive mass rural residents, including migrants, to towns or small cities near the villages should be implemented very prudently with more policies. Farmers driven to towns without land, but with some cash accumulated from migration, would actually consume more in the short-term as per the Government's plan. However, questions need to be asked regarding their employment after losing land, their public pension and medical service, or the access to public schools for the children of migrants. These issues, among many others, would make the policy makers think more earnestly. Migration and mass media are both contributing to rural development. Considering the strong effect of mass media exposure, policy makers may want to address the mass media in villages. The information and knowledge learnt from the media are important for the rural residents. Having access to more TV channels and Internet broadband would predict a significant increase in the knowledge of the rural residents,

not only on diet, but also on many other aspects such as agriculture technique, or prices of agriculture products.

6

Appendix

6.1 Variables and Measurement for Chapter 2

This appendix presents the variables and measurements for Chapter 2. The variables are listed in Table 6.1 and Table 6.2 (continued). These variables are the dependent and independent variables in the analysis. Individual, household and village controls are explained and the range of the values are presented in the table. Detailed summary statistics of the variables can be found in Table 2.2. Similarly, variables and measurement tables for Chapter 3 and Chapter 4 are also presented. The tables are organised and displayed in the same way as the following table. Without a specific need, the variables and measurements tables in the appendix will not be separately explained. As the analysis in the three chapters utilises the same data set but a different interested sample from different studies, certain variables may be the same, but the value and range may slightly alter.

Table 6.1: Variables and measurement for Chapter 2

	Variable	Description	Min	Max
Outcomes	agriculture participation	do you participate in agriculture activities?	0	1
	agriculture hours	how many hours per day do you spend on agriculture activities	0	14
	housework participation	do you participate in any housework activities?	0	1
	housework hours	how many hours per day do you spend on housework	0	16
	wagelabor participation	do you participate in wagelabor	0	1
	wagelabor hour	how many hours per day do you spend on wagelabor	0	12
Individual Controls	children age	the age of the children	6	18
	children education	the current level of education of the children	0	2
	children school enroll	do this children enroll in school?	0	1
Household Controls	family size	how many people live in the family?	1	13
	male household head age	how old is the father of the child?	22	60
	female household head age	how old is the mother of the child?	20	56
	grandmother living	does the grandmother live in the family?	0	1
	traffic vehicle	does this household own any traffic vehicle?	0	1
	farm machine	does this household own any farm machine?	0	1
	white goods	does this household own any white goods?	0	1
	information	does this household own any machine to receive information? Such as TV, phone	0	1
	chore tools	does this household own any chore tool, such as cook pot	0	1

Table continues on next page

Table 6.2: Variables and measurement for Chapter 2: continue

Village Controls	urbanisation index	urbanization index	22.45	103.07
	population density	community population density category	2	10
	diversity	diversity score	2.5	9
	economics	economic component score	0	10
	health	quality of health score	0	10
	housing	housing component score	0	10
	market	market component score	0	10
	social service	social services score	0	10
	transportation	transportation component score	0	10
	education	community education category	0.48	7.94
	modern markets	modern markets component score	0	10
	sanitation	sanitation score	0	10

This table presents the variables and description for Chapter 2.

6.2 Non-attritors and Attritors for Chapter 2

Attritors are the individuals who are only observed once, and thus discarded, while non-attritors are the ones that are observed at least twice in order to implement the fixed effects models.

Table 6.3: Compare attritors and non-attritors for Chapter 2

Year	Male		Female		
Variable	non-attritors		attritors		difference
	Mean	Std. Err.	Mean	Std. Err.	
Father migration	0.10	0.00	0.11	0.01	-0.01
Children age	12.28	0.04	12.43	0.10	-0.15
Children education	0.37	0.01	0.41	0.02	-0.04*
Children school enroll	0.88	0.00	0.87	0.01	0.01
Family size	4.42	0.01	4.59	0.03	-0.17***
Male household head age	39.67	0.07	39.85	0.15	-0.18
Female household head age	38.33	0.07	38.60	0.15	-0.27**
Grandmother living	0.19	0.00	0.20	0.01	-0.01
Agriculture par	0.06	0.00	0.06	0.01	0.00
Agriculture hours	0.21	0.01	0.23	0.05	-0.02
Housework par	0.22	0.00	0.24	0.01	-0.02**
Housework hours	0.18	0.01	0.17	0.02	0.01
Wage labor par	0.02	0.00	0.02	0.00	0.00
Wage labor hour	0.14	0.01	0.18	0.03	-0.04*
Traffic vehicle	0.84	0.00	0.80	0.01	0.04***
Farm machine	0.29	0.01	0.28	0.01	0.01
White goods	1.00	0.01	1.02	0.03	-0.02
Information	0.97	0.00	0.96	0.00	0.01**
Chore tools	1.82	0.01	1.79	0.02	-0.03
N	7,226		1,725		

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$. The attritors are those observations that are observed only once and the non-attritors are the observations that are observed at least twice. N is the number of observations.

6.3 Migration and Return in Chapter 3

Table 6.4 sets out the times of migration of individuals for each survey year. "Year" in the first column means the time period in which an individual is observed for the first time, while the numbers 0 to 6 are the number of times of migration is recorded for the individuals over their entire observed duration. In 1997, 5,436 were non-migrants, 1,005 individuals migrated once, and 14 people migrated 6 times. In 2000, amongst the new members joining the survey this year (2000), 1,209 were non-migrants, and 12 people migrated 5 times. From this table, a detailed picture can be ascertained regarding the data structure on migration and return.

Table 6.4: Migration times of individuals in Chapter 3

Year	0	1	2	3	4	5	6
1997	5,436	1,005	483	266	121	62	14
2000	1,209	244	104	49	32	12	
2004	654	119	58	30	10		
2006	480	187	89	31			
2009	737	196	54				
2011	1,804	187					

Migration times of 15-70 years old population. "Year" is the first time an individual is observed. "Migration times" is the total number of times individuals migrated over the entire observed period.

Table 6.5 presents the number of return migrants who migrate again in the following wave. The rates in the table are calculated as follows: $Rate = \frac{return_{t,m}}{return_t}$ with $t \in \{2000, 2004, 2006, 2009\}$; m is an indicator for migration in wave $t+1$. The returnees' migration rates range between 20% to 60% over the years. This evidence shows that the rural to urban migration in China is temporary, with a significant percentage of return migrants migrating again after returning. These return migrants have experienced an

urban life style and a world beyond their rural villages. As such, they bring changes to their home villages and non-migrants, including adjustments in political ideas, social norms, and especially on the consumption behaviour this paper focuses on.

Table 6.5: Returnees who migrate again in Chapter 3

Year	Male		Female	
	Returnees	Rate	Returnees	Rate
2000	65	0.60	14	0.27
2004	86	0.43	21	0.24
2006	110	0.32	44	0.29
2009	141	0.24	48	0.21

Returnees who migrate again in the following panel

Table 6.6: Survey time in Chapter 3

Survey Month		
Month	Frequency	Rate
JAN	2,860	1.85
JUL	713	0.46
AUG	9,791	6.35
SEP	48,149	31.22
OCT	56,735	36.79
NOV	27,655	17.93
DEC	8,315	5.39

Survey time of the whole survey.

6.4 IV on Current Migration and Return in Chapter 3

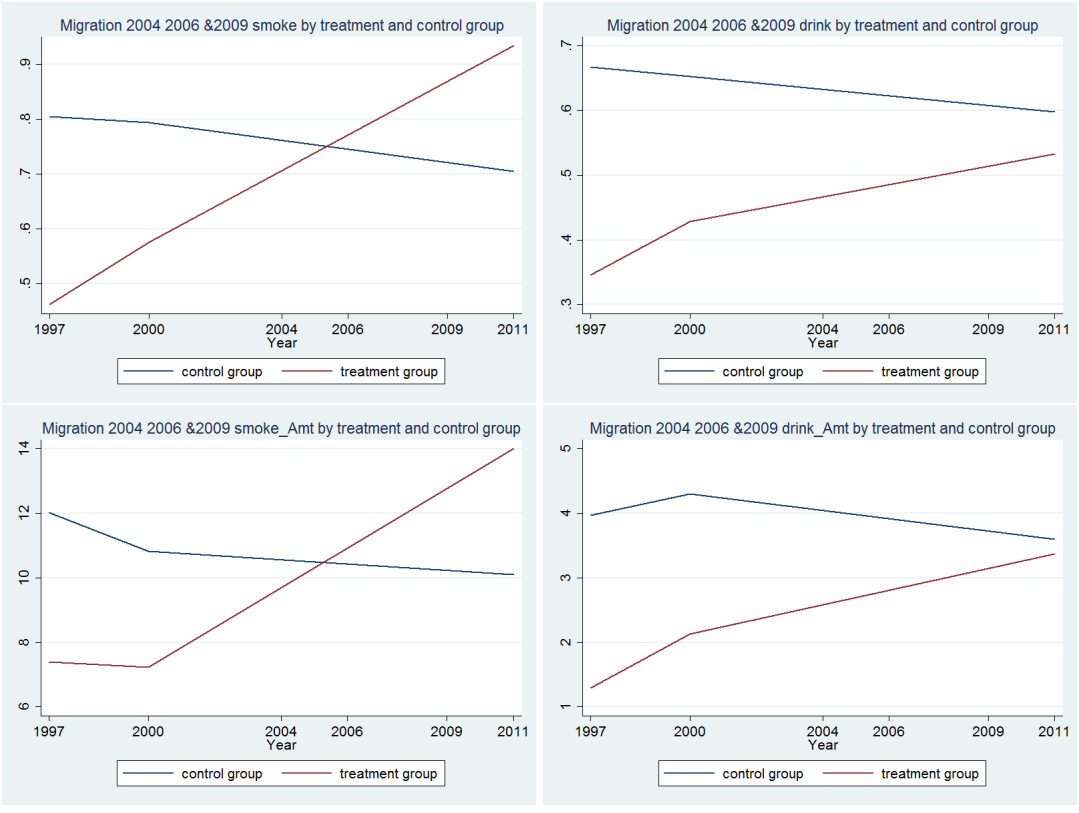
Regression results in table 6.7 illustrate the correlation of IV with migration and return migration status. These regression models are bivariate analysis with the instrumental variable as an independent variable and migration or return migration as dependent variable. Robust standard errors are applied wherever available to the highest level of robust. In panel Logit FE model, the bootstrap robust standard errors are applied. The results demonstrate the strong correlation between the IV and migration/return migration.

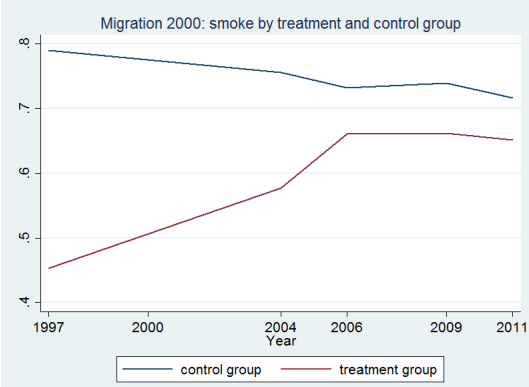
Table 6.7: Regression model IV on migration and return for Chapter 3

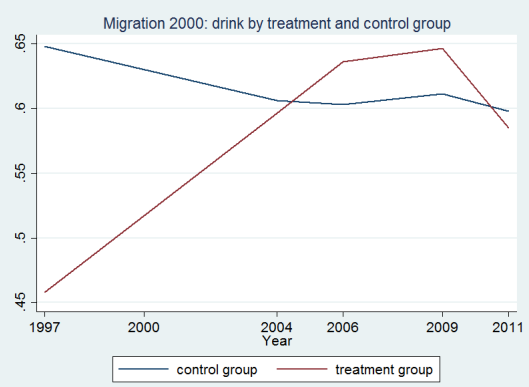
Dependent Variable	Migration	Return
Pooled OLS	0.10***	0.02***
	(0.00)	(0.00)
Linear FE	0.44***	0.40***
	(0.02)	(0.02)
Panel Probit	0.85***	0.73***
	(0.05)	(0.07)
Panel Logit FE	3.27***	12.01***
	(0.21)	(0.89)

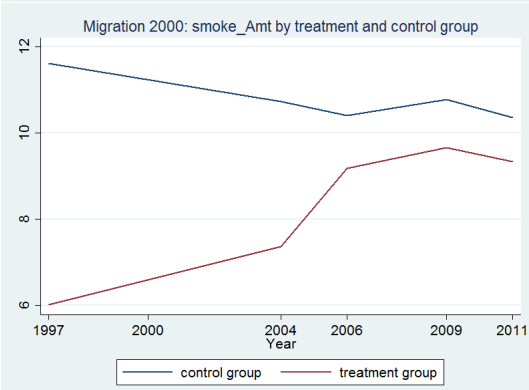
Standard errors in parentheses.

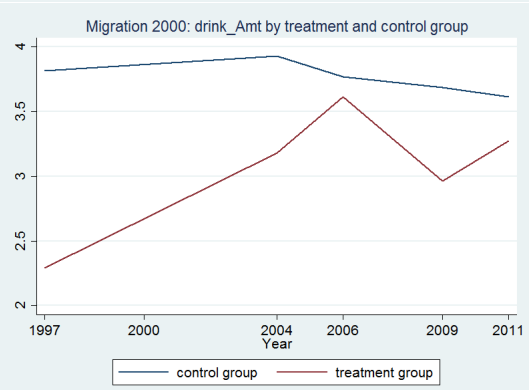
6.5 DiD Graphs for Separated Panels in Chapter 3

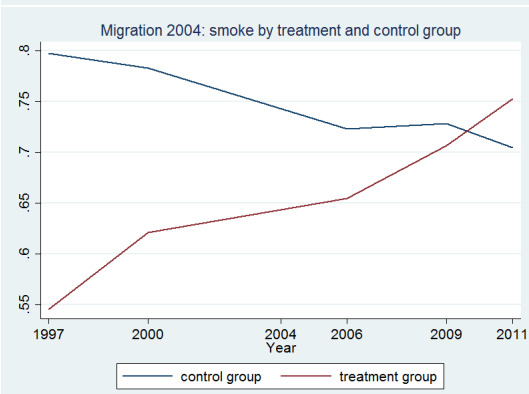


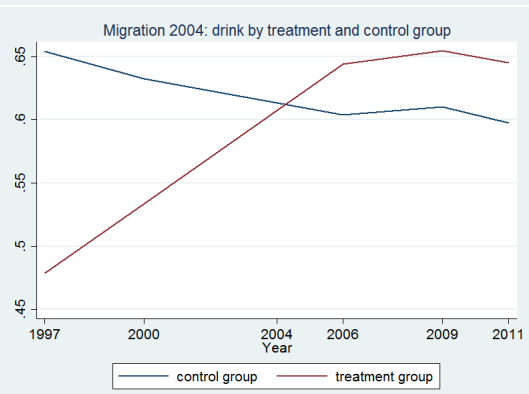

Migration 2000: smoke by treatment and control group
.4
.5
.6
.7
.8
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

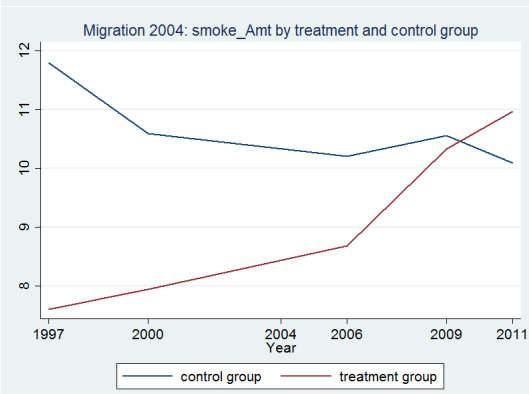

Migration 2000: drink by treatment and control group
.45
.5
.55
.6
.65
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

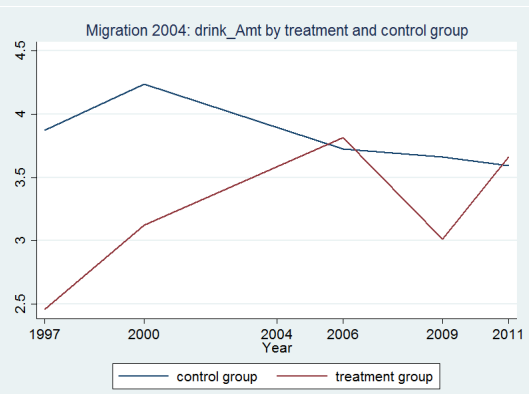

Migration 2000: smoke_Amt by treatment and control group
6
8
10
12
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

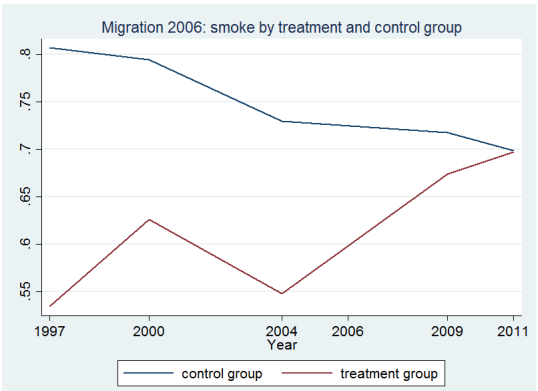

Migration 2000: drink_Amt by treatment and control group
2
2.5
3
3.5
4
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

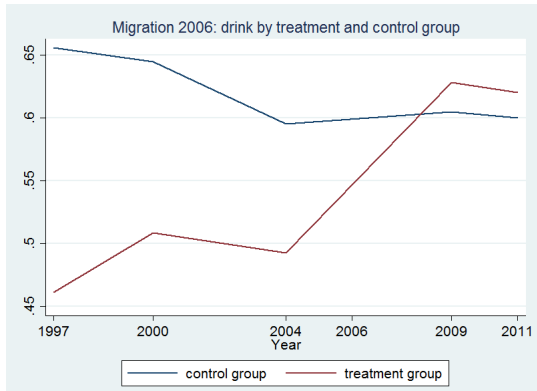

Migration 2004: smoke by treatment and control group
.55
.6
.65
.7
.75
.8
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

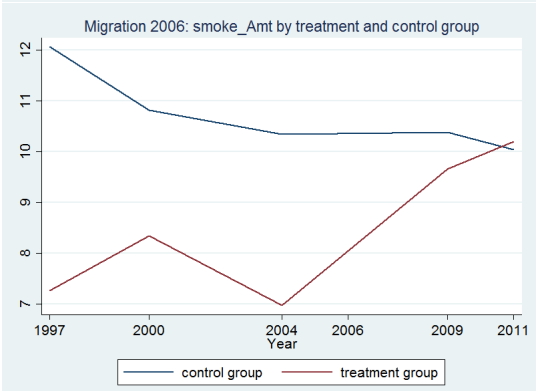

Migration 2004: drink by treatment and control group
.45
.5
.55
.6
.65
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

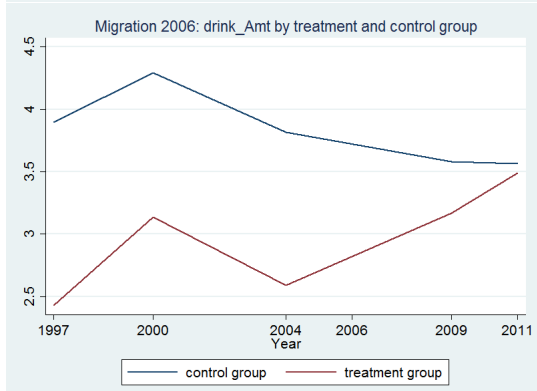

Migration 2004: smoke_Amt by treatment and control group
8
9
10
11
12
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

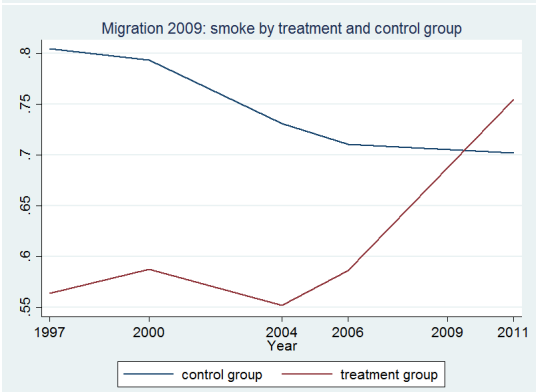

Migration 2004: drink_Amt by treatment and control group
2.5
3
3.5
4
4.5
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

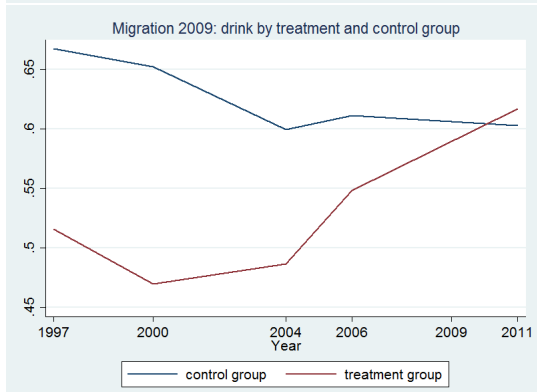

Migration 2006: smoke by treatment and control group
.55
.6
.65
.7
.75
.8
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

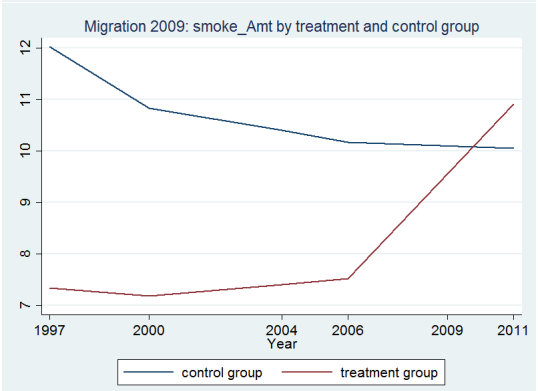

Migration 2006: drink by treatment and control group
.45
.5
.55
.6
.65
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

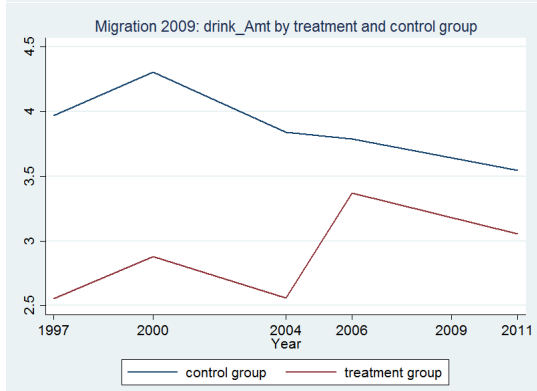

Migration 2006: smoke_Amt by treatment and control group
7
8
9
10
11
12
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

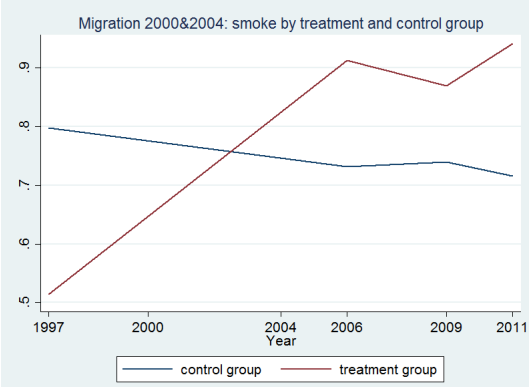

Migration 2006: drink_Amt by treatment and control group
2.5
3
3.5
4
4.5
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

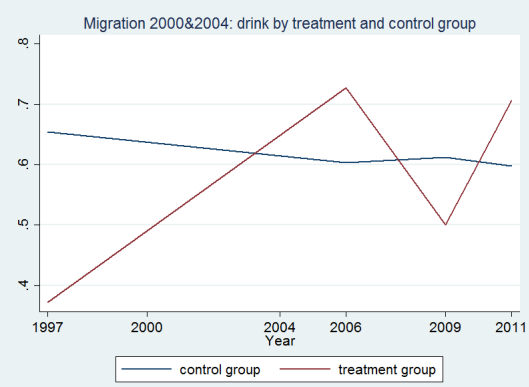

Migration 2009: smoke by treatment and control group
.55
.6
.65
.7
.75
.8
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

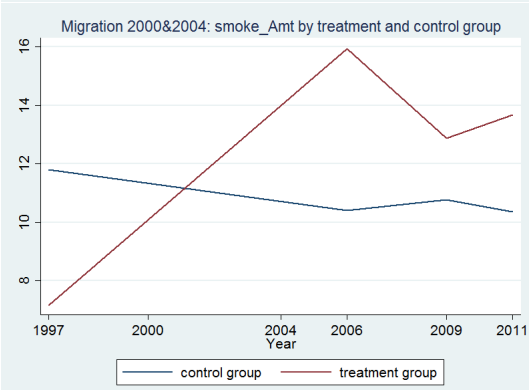

Migration 2009: drink by treatment and control group
.45
.5
.55
.6
.65
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

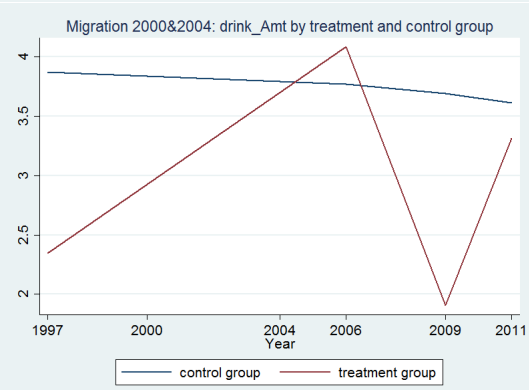

Migration 2009: smoke_Amt by treatment and control group
7
8
9
10
11
12
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

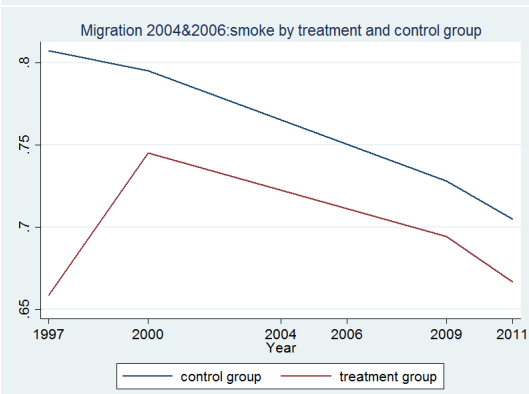

Migration 2009: drink_Amt by treatment and control group
2.5
3
3.5
4
4.5
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

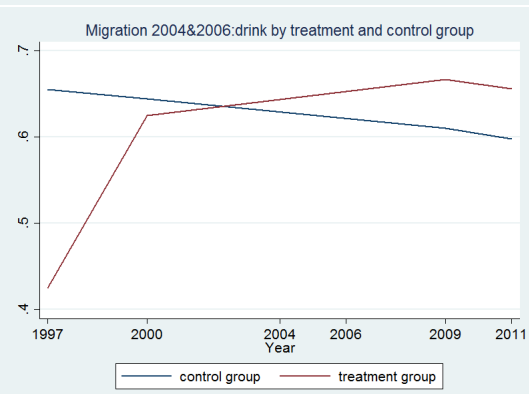

Migration 2000&2004: smoke by treatment and control group
.5
.6
.7
.8
.9
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

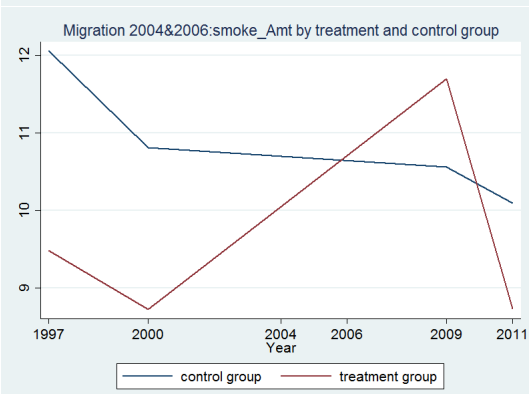

Migration 2000&2004: drink by treatment and control group
.4
.5
.6
.7
.8
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

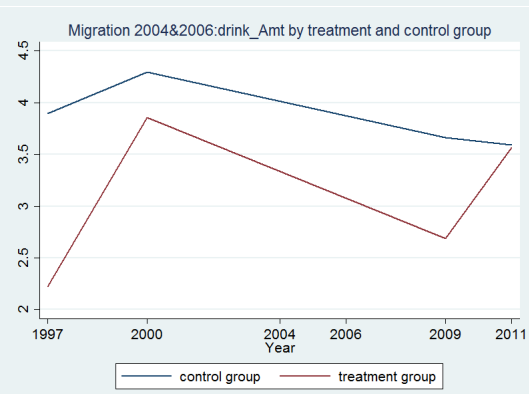

Migration 2000&2004: smoke_Amt by treatment and control group
8
10
12
14
16
1997
2000
2004
2006
2009
2011
Year
control group
treatment group

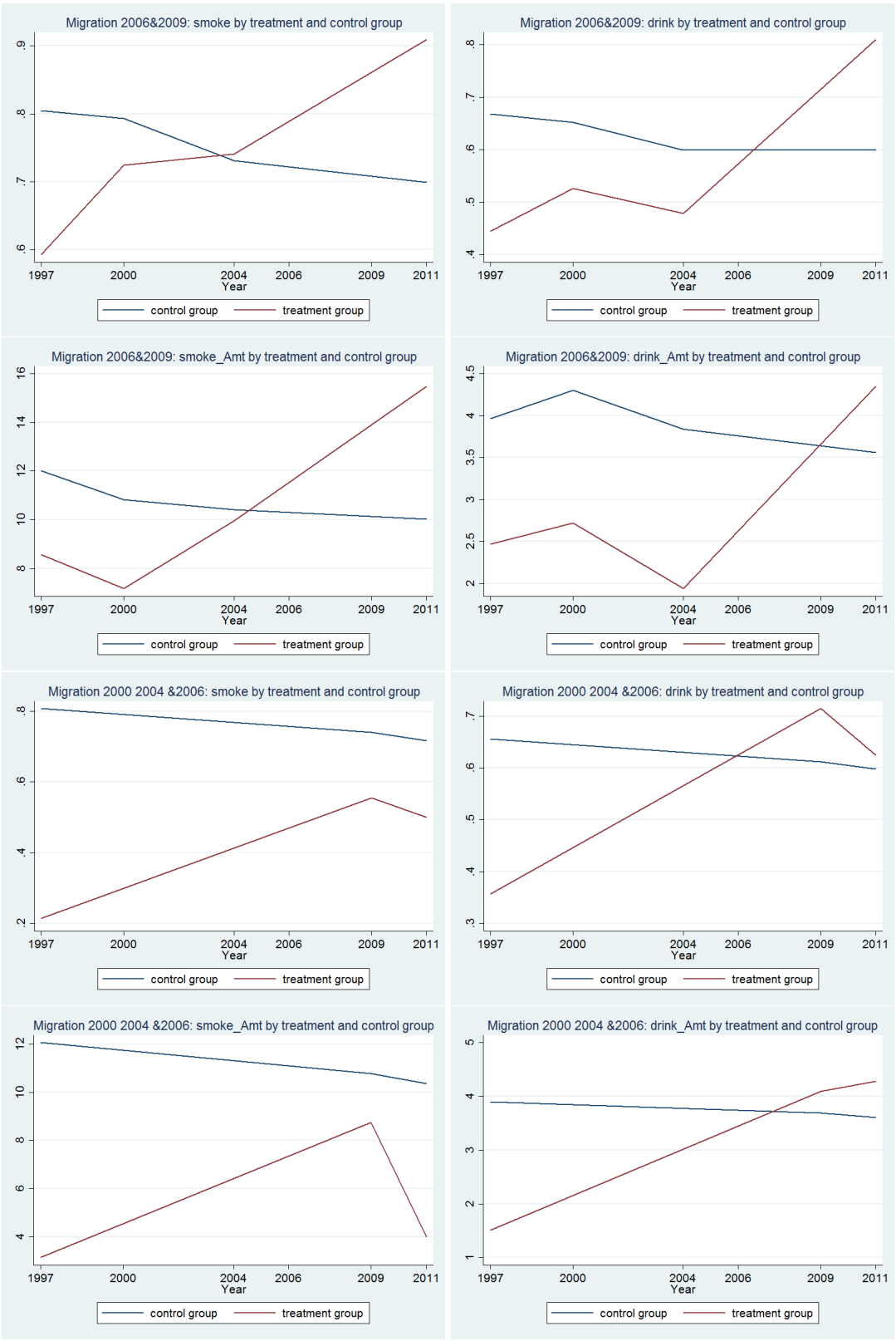

Migration 2000&2004: drink_Amt by treatment and control group
2
2.5
3
3.5
4
1997
2000
2004
2006
2009
2011
Year
control group
treatment group


Migration 2004&2006:smoke by treatment and control group
.65
.7
.75
.8
1997
2000
2004
2006
2009
2011
Year
control group
treatment group


Migration 2004&2006:drink by treatment and control group
.4
.5
.6
.7
1997
2000
2004
2006
2009
2011
Year
control group
treatment group


Migration 2004&2006:smoke_Amt by treatment and control group
9
10
11
12
1997
2000
2004
2006
2009
2011
Year
control group
treatment group


Migration 2004&2006:drink_Amt by treatment and control group
2
2.5
3
3.5
4
4.5
1997
2000
2004
2006
2009
2011
Year
control group
treatment group



CHNS 1997 to 2011, 15 to 70 years old male rural residents.

6.6 Variables and Measurement for Chapter 3

Table 6.8: Variables and measurement for Chapter 3

	Variables	Description	Min	Max
Smoking	smoke	does this individual smoke or not?	0	1
	smoke amount	how many cigarettes do you smoke per day?	0	20
Drinking	drink	does this individual drink or not?	0	1
	drink amount	how much alcohol do you drink each week?	0	10
Individual controls	age	calculated age in years to 0 decimal points	15	70
	age square	square of age	225	4,900
	marriage	marriage status of individual	0	1
	education	highest level of education attained	0	2
	sick	has the individual been sick or injured in last 4 weeks?	0	1
	medical insurance	do you have medical insurance of any kinds?	0	1
Household controls	own wash machine	does any household member own washing machine?	0	1
	own phone	does household own phone?	0	1
	fuel type	type of fuel the household normally used for cooking	0	1

Table continues on next page

Table 6.9: Variables and measurement for Chapter 3: continue

Village controls	urbanization index	urbanization index	22.45	103.07
	population density	community population density category	2	10
	diversity	diversity score	2.5	9.5
	economics	economic component score	0	10
	health	quality of health score	0	10
	housing	housing component score	0	10
	market	market component score	0	10
	social services	social services score	0	10
	transportation	transportation component score	0	10
	education	community education category	0.48	7.94
	modern markets	modern markets component score	0	10
	sanitation	sanitation score	0	10

This table shows the variables and description in Chapter 3.

6.7 Male Smoking and Drinking Behaviours in Rural and Urban China for Chapter 3

Table 6.10 compares the smoking and drinking behaviour of rural and urban males in China and depicts that urban males are likely to smoke and drink more.

Table 6.10: Male smoking and drinking behaviours in rural and urban China for Chapter 3

	Urban		Rural		Diff
	Mean	Std. Err.	Mean	Std. Err	
Smoke	0.707	0.005	0.676	0.003	0.031***
Smoke Amount	11.060	0.089	9.914	0.062	1.146***
Drink	0.604	0.005	0.554	0.004	0.050***
Drink Amount	3.438	0.041	3.392	0.029	0.046*

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$

6.8 Compare Migrants and Returnees for Chapter 3

Table 6.11 compares the current migrants and return migrants on the basis of certain observables.

Table 6.11: Returnees and migrants for Chapter 3

	Migrants		Returnees		Diff
	Mean	Std. Err.	Mean	Std. Err.	
Age	34.76	0.15	34.50	0.18	0.26
Marriage	0.61	0.01	0.60	0.01	0.01
Education	0.81	0.01	0.83	0.01	-0.02*
Height	166.36	0.12	166.45	0.10	-0.09
Weight	61.95	0.18	62.17	0.16	-0.22
Med Insurance	0.16	0.01	0.34	0.01	-0.18***
Energy	0.64	0.01	0.65	0.01	-0.01
Wash Machine	0.50	0.01	0.52	0.01	-0.02**
Phone	0.45	0.01	0.44	0.01	0.01
Urbanisation index	51.97	0.22	51.67	0.22	0.30
Population density	5.52	0.02	5.62	0.02	-0.10
Diversity	4.42	0.01	4.41	0.01	0.01
Economy	4.70	0.04	4.62	0.04	0.08*
Health	4.55	0.03	4.48	0.03	0.07**
House	5.54	0.30	5.54	0.31	0.00
Market	3.77	0.05	3.71	0.05	0.06
Social Service	2.03	0.03	2.06	0.03	-0.03
Transportation	5.09	0.03	4.98	0.03	0.11**
Education	2.62	0.01	2.60	0.01	0.02*
Modern Market	3.37	0.04	3.36	0.04	0.01
Sanitation	4.99	0.04	4.96	0.04	0.03

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$, please note that as current migrants do not appear for the survey interview, many other characteristics of the current migrants are not available in the data, for instance, their smoking and drinking behaviour.

6.9 Variables and Measurement for Chapter 4

Table 6.12: Variables and explanation for Chapter 4

	Variables	Description	Min	Max
Diet Knowledge	$diet\_i, t$	diet knowledge of individual PCA	-6.87	1.66
Village migration	mig_v, t	$mig\_rate_{v,t}$ village migration rate	0	0.95
		$mig\_den_{v,t}$ proportion migrant household	0	0.40
Mass media exposure	$mass\_i, t$	$mass\_tv_{i,t}$ dummy variable, individual participation in watching TV	0	1
		$mass\_tv\_time_{i,t}$ individual time spent on TV, hours per week	0	41.50
		$mass\_int_{i,t}$ dummy variable, individual use Internet or not	0	1
		$mass\_int\_time_{i,t}$ individual time spend on Internet, hours per week	0	50.05
Knowledge peer effect	$peer\_i, t$	neighbourhood diet knowledge to capture peer effect	-4.21	-0.03
Education level	education-middle	dummy variable, 1 for middle level education	0	1
	education-high	dummy variable, 1 for higher level education	0	1
Individual controls	age	calculated age in years to 0 decimal points	20	55
	gender	gender of individual	1	2
	nationality	nationality of individual	0	1
	education	highest level of education attained, uses the education category in regressions	0	2

table continues next page

Table 6.13: Variables and explanation for Chapter 4: continue

Household controls	household income	total net household income	-13.00	13.00
	pay water	does household pay for drinking water?	0	1
	toilet facilities	type of toilet facilities in household	0	1
	own bike	does any household member own bicycle?	0	1
	own wash machine	does any household member own washing machine?	0	1
	own fridge	does any household member own refrigerator?	0	1
	own elc-fan	does any household member own electric fan?	0	1
	own ovens	does household own microwave?	0	1
	own elc-pot	does household own electric cooking pot?	0	1
	own phone	does household own phone?	0	1
	fuel type	type of fuel the household normally used for cooking	0	1
Village controls	observed Population	population in the village observed in the data	3	172
	number of households	number of households in the village observed in the data	1	30
	urbanization index	urbanization index	23.39	103.07
	population density	community population density category	2	9.5
	diversity	diversity score	2.5	9.5
	economics	economic component score	0	10
	health	quality of health score	0.8	10
	housing	housing component score	0.97	10
	market	market component score	0	10
	social services	social services score	0	10
	transportation	transportation component score	0	10
	education	community education category	0.48	7.94
	modern markets	modern markets component score	0	10
	sanitation	sanitation score	0	10

6.10 Diet knowledge Questionnaire in Chapter 4

Table 6.14: Diet knowledge questions list for Chapter 4

Question number	Statement
1	Choosing a diet with lots of fresh fruits and vegetables is good for one's health.
2	Eating a lot of sugar is good for one's health.
3	Eating a variety of foods is good for one's health.
4	Choosing a diet high in fat is good for one's health.
5	Choosing a diet with a log of staple food [rice and rice products and wheat and wheat products] is not good for one's health.
6	Consuming a lot of animal products daily [fish, poultry, eggs and lean meat] is good for one's health.
7	Reducing the amount of fatty meat and animal fat in the diet is good for one's health.
8	The heavier the body weight, the healthier the individual is.
9	Consuming milk and dairy products is good for one's health.
10	Consuming beans and bean products is good for one's health.

The interviewers ask the respondent if he or she strongly agree, somewhat agree, neutral, somewhat disagree or strongly disagree with each statement above? Question 8 is included as it reflects some knowledge about diet and health, even though, it is not a direct statement on a particular type of food.

6.11 Diet Knowledge and Behaviour for Chapter 4

This appendix presents the correlation between diet knowledge and health behaviour (smoking and drinking) to show that diet knowledge studied in Chapter 4 indeed have some correlation with the health behaviour, even though the study outcome is not directly dietary behaviour.

Table 6.15: Correlation matrices: Diet knowledge and behaviour for Chapter 4

	Diet Knowledge	Smoke	Smoke Amount	Drink	Drink Amount
Smoke	-0.012**	1.000			
Smoke Amount	-0.016*	0.912***	1.000		
Drink	-0.024***	0.500***	0.498***	1.000	
Drink Amount	-0.003*	0.486***	0.511***	0.865***	1.000

$^{*}p < 0.1$, $^{**}p < 0.05$ and $^{***}p < 0.01$

Table 6.16: Regression results: Diet knowledge and behaviour for Chapter 4

	smoke		smoke amount
Diet Knowledge	-0.002***	Diet Knowledge	-0.006**
	(0.001)		(0.003)
	drink		drink amount
Diet Knowledge	-0.006***	Diet Knowledge	-0.016**
	(0.002)		(0.008)

Standard errors in parentheses.

References

Acosta, P. (2011). School attendance, child labour, and remittances from international migration in El Salvador. *Journal of Development Studies*, 47(6):913–936.

Aghion, P. and Jaravel, X. (2015). Knowledge spillovers, innovation and growth. *The Economic Journal*, 125(583):533–573.

Amuedo-Dorantes, C. and Pozo, S. (2006). Migration, remittances, and male and female employment patterns. *The American economic review*, pages 222–226.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Antman, F. M. (2010a). Adult child migration and the health of elderly parents left behind in Mexico. *American Economic Review*, 100(2):205–08.

Antman, F. M. (2010b). International migration, spousal control, and gender discrimination in the allocation of household resources. *University of Colorado at Boulder Department of Economics Working Paper*, (10-15).

Antman, F. M. (2011a). The intergenerational effects of paternal migration on school-

ing and work: What can we learn from children's time allocations? *Journal of Development Economics*, 96(2):200–208.

Antman, F. M. (2011b). International migration and gender discrimination among children left behind. *The American economic review*, 101(3):645.

Antman, F. M. (2012a). Elderly care and intrafamily resource allocation when children migrate. *Journal of Human Resources*, 47(2):331–363.

Antman, F. M. (2012b). Gender, educational attainment, and the impact of parental migration on children left behind. *Journal of Population Economics*, 25(4):1187–1214.

Antman, F. M. (2013). The impact of migration on family left behind. *International Handbook on the Economics of Migration*, page 293.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144).

Barro, R. and Sala-i Martin, X. (1997). Technological diffusion, convergence, and growth. *Journal of Economic Growth*, 2(1):1–26.

Batista, C. and Vicente, P. C. (2011). Do migrants improve governance at home? evidence from a voting experiment. *World Bank Economic Review*, 25(1):77–104.

Bertoli, S. and Marchetta, F. (2015). Bringing it all back home return migration and fertility choices. *World Development*, 65(0):27 – 40. Migration and Development.

Binzel, C. and Assaad, R. (2011). Egyptian men working abroad: Labour supply responses by the women left behind. *Labour Economics*, 18:98–114.

Blaylock, J., Smallwood, D., Kassel, K., Variyam, J., and Aldrich, L. (1999). Economics, food choices, and nutrition. *Food Policy*, 24(23):269 – 286.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393.

Bursztyn, L. and Cantoni, D. (2015). A tear in the iron curtain: The impact of western television on consumption behavior. *Review of Economics and Statistics*.

Cabegin, E. (2006). The effect of Filipino overseas migration on the non-migrant spouse's market participation and labor supply behavior. *IZA Discussion Paper*, 2240.

Case, A. and Paxson, C. (2006). Stature and status: Height, ability, and labor market outcomes. Working Paper 12466, National Bureau of Economic Research.

Chan, K. W. (2013). China: internal migration. *The encyclopedia of global human migration*.

Chang, H., Dong, X., and MacPhail, F. (2011). Labor migration and time use patterns of the left-behind children and elderly in rural China. *World Development*, 39(12):2199–2210.

Chen, J. J. (2006). Migration and imperfect monitoring: implications for intra-household allocation. *The American economic review*, pages 227–231.

Chen, J. J. (2013). Identifying non-cooperative behavior among spouses: child outcomes in migrant-sending households. *Journal of Development Economics*, 100(1):1–18.

Chen, X., Huang, Q., Rozelle, S., Shi, Y., and Zhang, L. (2009a). Effect of migration on children's educational performance in rural China. *Comparative Economic Studies*, 51(3):323–343.

Chen, X., Stanton, B., Li, X., Fang, X., Lin, D., and Xiong, Q. (2009b). A comparison of health-risk behaviors of rural migrants with rural residents and urban residents in China. *American Journal of Health Behavior*, 33(1):15–25.

Chen, Y. and Feng, S. (2013). Access to public schools and the education of migrant children in China. *China Economic Review*, 26:75–88.

Chen, Y., Jin, G. Z., and Yue, Y. (2010). Peer migration in China. Technical report, National Bureau of Economic Research.

Chew, F., Palmer, S., Slonska, Z., and Subbiah, K. (2002). Enhancing health knowledge, health beliefs, and health behavior in Poland through a health promoting television program series. *Journal of Health Communication*, 7(3):179–196.

Chiang, Y., Hannum, E., and Kao, G. (2012). Who goes, who stays, and who studies? gender, migration, and educational decisions among rural youth in China. *International Journal of Chinese Education*, 1(1):106–131.

Clemens, M. A., Özden, C., and Rapoport, H. (2014). Migration and development research is moving far beyond remittances. *World Development*, 64(0):121 – 124.

Correia, S. (2014). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. Statistical Software Components, Boston College Department of Economics.

Curtis, C. C., Lugauer, S., and Mark, N. C. (2011). Demographic patterns and household saving in China. Technical Report 16828, National Bureau of Economic Research.

Dahl, G. and DellaVigna, S. (2009). Does movie violence increase violent crime? *The Quarterly Journal of Economics*, 124(2):677–734.

De Brauw, A. and Giles, J. (2008). Migrant opportunity and the educational attainment of youth in rural China. *World Bank Policy Research Working Paper Series*.

De Brauw, A. and Giles, J. (2012). Migrant labor markets and the welfare of rural households in the developing world: Evidence from China. *IZA Discussion Paper*.

DellaVigna, S. and Ferrara, E. L. (2015). Economic and social impacts of the media. Working Paper 21360, National Bureau of Economic Research.

DellaVigna, S. and Kaplan, E. (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, 122:1187–1234.

Démurger, S. and Li, S. (2013). Migration, remittances, and rural employment patterns: Evidence from China. *Labor Market Issues in China*, page 31.

Démurger, S. and Xu, H. (2011). Return migrants: The rise of new entrepreneurs in rural China. *World Development*, 39(10):1847 – 1861.

Démurger, S. and Xu, H. (2013). Left-behind children and return decisions of rural

migrants in China. *IZA Discussion Paper*, (7727).

Durbin, J. (1954). Errors in variables. *Review of the International Statistical Institute*, 22:23–32.

Dustmann, C. and Görlach, J.-S. (2016). The economics of temporary migrations. *Journal of Economic Literature*, 54(1):98–136.

Edwards, A. C. and Ureta, M. (2003). International migration, remittances, and schooling: evidence from El Salvador. *Journal of development economics*, 72(2):429–461.

Fan, C. C. (2003). Rural-urban migration and gender division of labor in transitional China. *International Journal of Urban and Regional Research*, 27(1):24–47.

Fei, L. (2007). Rural-urban migration in China: Recent trends and future challenges. In *International Conference on Contemporary China Studies*.

Frijters, P., Kong, T., and Meng, X. (2011). Migrant entrepreneurs and credit constraints under labour market discrimination. *IZA Discussion Paper*.

Gentzkow, M. (2006). Television and voter turnout. *The Quarterly Journal of Economics*, 121(3):pp. 931–972.

Gentzkow, M. and Shapiro, J. M. (2008). Preschool television viewing and adolescent test scores: Historical evidence from the coleman study. *The Quarterly Journal of Economics*, 123(1):pp. 279–323.

Gibson, J., McKenzie, D., and Stillman, S. (2011). What happens to diet and child health when migration splits households? Evidence from a migration lottery program. *Food Policy*, 36(1):7–15.

Giles, J. and Mu, R. (2007). Elderly parent health and the migration decisions of adult children: Evidence from rural China. *Demography*, 44(2):265–288.

Giles, J., Wang, D., and Zhao, C. (2010). Can Chinas rural elderly count on support from adult children? implications of rural-to-urban migration. *Journal of Population Ageing*, 3(3-4):183–204.

Giulietti, C., Wahba, J., and Zimmermann, K. F. (2013). Entrepreneurship of the left-behind. *Research in Labor Economics*, 37:65–92.

Goldberg, J. P. (1992). Nutrition and health communication: The message and the media over half a century. *Nutrition Reviews*, 50(3):71–77.

Görlich, D., Omar Mahmoud, T., and Trebesch, C. (2007). Explaining labour market inactivity in migrant-sending families: housework, hammock, or higher education? Technical Report 1391, Kieler Arbeitspapiere.

Grigorian, D. A. and Melkonyan, T. A. (2011). Destined to receive: The impact of remittances on household decisions in Armenia. *Review of Development Economics*, 15(1):139–153.

Guo, M., Aranda, M. P., and Silverstein, M. (2009). The impact of out-migration on the inter-generational support and psychological wellbeing of older adults in rural China. *Ageing and Society*, 29(07):1085–1104.

Hadi, A. (2001). International migration and the change of women's position among the left-behind in rural Bangladesh. *International Journal of Population Geography*, 7(1):53–61.

Handbury, J., Rahkovsky, I., and Schnell, M. (2015). What drives nutritional disparities? retail access and food purchases across the socioeconomic spectrum. Working Paper 21126, National Bureau of Economic Research.

Hanson, G. H. (2007). Emigration, remittances and labor force participation in Mexico. *Working Paper ITD= Documento de Trabajo ITD; 28*.

Harper, B. (2000). Beauty, stature and the labour market: A British cohort study. *Oxford Bulletin of Economics and Statistics*, 62:771–800.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica*, 46(6):1251–71.

Head, R., Murray, J., Sarrassat, S., Snell, W., Meda, N., Ouedraogo, M., Deboise, L., and Cousens, S. (2015). Can mass media interventions reduce child mortality? *The*

Lancet, 386(9988):97–100.

Hildebrandt, N., McKenzie, D. J., Esquivel, G., and Schargrodsky, E. (2005). The effects of migration on child health in Mexico. *Economia*, pages 257–289.

H.Nagler, R. (2014). Adverse outcomes associated with media exposure to contradictory nutrition messages. *Journal of Health Communication*, 19(1):24–40.

Hornung, E. (2014). Immigration and the diffusion of technology: The Huguenot diaspora in Prussia. *American Economic Review*, 104(1):84–122.

Jaffe, A. B. (1986). Technological opportunity and spillovers of R & D: Evidence from firms' patents, profits, and market value. *The American Economic Review*, 76(5):984–1001.

Jensen, R. and Oster, E. (2009). The power of TV: Cable television and women's status in India. *The Quarterly Journal of Economics*, 124(3):1057–1094.

Jiang, H., Room, R., and Hao, W. (2015). Alcohol and related health issues in China: action needed. *The Lancet Global Health*, 3(4):190 – 191.

Johnston, D. W., Lordan, G., Shields, M. A., and Suziedelyte, A. (2015). Education and health knowledge: Evidence from UK compulsory schooling reform. *Social Science & Medicine*, 127:92 – 100.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

Kenkel, D. S. (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2):pp. 287–305.

Kerr, W. R. (2008). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics*, 90(3):518–537.

Kim, N. (2007). The impact of remittances on labor supply: The case of Jamaica. *World Bank Policy Research Working Paper*, (4120).

Knight, J., Deng, Q., and Li, S. (2011). The puzzle of migrant labour shortage and

rural labour surplus in China. *China Economic Review*, 22(4):585–600.

Knight, J. and Gunatilaka, R. (2010). Great expectations? the subjective well-being of rural–urban migrants in China. *World Development*, 38(1):113–124.

Knight, J. and Song, L. (2003). Chinese peasant choices: migration, rural industry or farming. *Oxford Development Studies*, 31(2):123–148.

Lancet (2014). A step change for tobacco control in China? *The Lancet*, 384(9959):2000.

Lee, M.-H. (2011). Migration and children's welfare in China: The schooling and health of children left behind. *The Journal of Developing Areas*, pages 165–182.

Leng, L. and Albert, P. (2010). Parental migration and child development in China(working paper). *Gansu Survey of Children and Families*.

Li, Y., Jiang, Y., Zhang, M., Yin, P., Wu, F., and Zhao, W. (2011). Drinking behaviour among men and women in China: the 2007 China chronic disease and risk factor surveillance. *Addiction*, 106(11):1946–1956.

Lokshin, M. and Glinskaya, E. (2009). The effect of male migration on employment patterns of women in Nepal. *The World Bank Economic Review*.

Lu, Y. (2012). Education of children left behind in rural China. *Journal of Marriage and Family*, 74(2):328–341.

Ma, Z. (2001). Urban labour-force experience as a determinant of rural occupation change: evidence from recent urban - rural return migration in China. *Environment and Planning A*, 33(2):237–255.

Ma, Z. (2002). Social-capital mobilization and income returns to entrepreneurship: the case of return migration in rural China. *Environment and Planning A*, 34(10):1763–1784.

Mansuri, G. (2006). Migration, sex bias, and child growth in rural Pakistan. *World Bank policy research working paper*, (3946).

McCormick, B. and Wahba, J. (2001). Overseas work experience, savings and entrepreneurship amongst return migrants to ldcs. *Scottish Journal of Political Economy*, 48(2):164–178.

McKenzie, D. and Rapoport, H. (2007). Network effects and the dynamics of migration and inequality: theory and evidence from Mexico. *Journal of development Economics*, 84(1):1–24.

Meng, X. (2012). Labor market outcomes and reforms in China. *The Journal of Economic Perspectives*, 26(4):pp. 75–101.

Meng, X. and Zhang, J. (2001). The two-tier labor market in urban China: Occupational segregation and wage differentials between urban residents and rural migrants in shanghai. *Journal of Comparative Economics*, 29(3):485 – 504.

Meyerhoefer, C. D. and Chen, C. (2011). The effect of parental labor migration on childrens educational progress in rural China. *Review of Economics of the Household*, 9(3):379–396.

Mu, R. and Van de Walle, D. (2011). Left behind to farm? women's labor re-allocation in rural China. *Labour Economics*, 18:83–97.

Parodi, G., Pastore, F., Giulietti, C., Ning, G., and Zimmermann, K. F. (2012). Self-employment of rural-to-urban migrants in China. *International Journal of Manpower*, 33(1):96–117.

Poncet, S. and Zhu, N. (2005). Country people moving to the cities: the migratory dynamic. *China perspectives*, (62).

Prat, A. and Stromberg, D. (2013). The political economy of mass media. *Advances in Economics and Econometrics: Volume 2, Applied Economics: Tenth World Congress*, 50:135.

Robert, J. and Oliver, H. (2010). Health and wages: Panel data estimates considering selection and endogeneity. *Journal of Human Resources*, 45(2):364–406.

Roberto, C. A., Swinburn, B., Hawkes, C., Huang, T. T.-K., Costa, S. A., Ashe, M.,

Zwicker, L., Cawley, J. H., and Brownell, K. D. (2015). Patchy progress on obesity prevention: emerging examples, entrenched barriers, and new thinking. *The Lancet*, 385(9985):2400 – 2409.

Rodriguez, E. R. and Tiongson, E. R. (2001). Temporary migration overseas and household labor supply: evidence from urban Philippines. *International Migration Review*, 35(3):709–725.

Rossi, A. (2008). The impact of migration on children in developing countries. In *Unpublished manuscript prepared for the Youth Migration Conference*, pages 24–26.

Rozelle, S., Taylor, J. E., and DeBrauw, A. (1999). Migration, remittances, and agricultural productivity in China. *The American Economic Review*, 89(2):287–291.

Sahn, D. E. and Alderman, H. (1988). The effects of human capital on wages, and the determinants of labor supply in a developing country. *Journal of Development Economics*, 29(2):157 – 183.

Shojaeizadeh, D. and Gashtaee, M. (2015). Assessing the relationship between knowledge attitude and healthy behaviour among menopaused women in Tehran in 2000. *Iranian Journal of Public Health*, 31(1-2):19–20.

Song, Z., Storesletten, K., and Zilibotti, F. (2011). Growing like China. *American Economic Review*, 101(1):196–233.

Spolaore, E. and Wacziarg, R. (2009). The diffusion of development. *The Quarterly Journal of Economics*, 124(2):469–529.

Spronk, I., Kullen, C., Burdon, C., and O'Connor, H. (2014). Relationship between nutrition knowledge and dietary intake. *British Journal of Nutrition*, 111(10):1713–1726.

Strauss, J. and Thomas, D. (1998). Health, nutrition, and economic development. *Journal of Economic Literature*, 36(2):766–817.

Sun, M. and Fan, C. C. (2011). China's permanent and temporary migrants: differentials and changes, 1990–2000. *The Professional Geographer*, 63(1):92–112.

Thomas, D. and Strauss, J. (1997). Health and wages: Evidence on men and women in urban Brazil. *Journal of Econometrics*, 77(1):159 – 185.

Tuccio, M. and Wahba, J. (2015). Can I have permission to leave the house? return migration and the transfer of gender norms. IZA Discussion Papers 9216, Institute for the Study of Labor (IZA).

Vogl, T. S. (2014). Height, skills, and labor market outcomes in Mexico. *Journal of Development Economics*, 107(0):84 – 96.

Wahba, J. (2014). Return migration and economic development. In Lucas, R. E., editor, *International Handbook on Migration and Economic Development*. Edward Elgar Publishing, Inc., Cheltenham, UK.

Wahba, J. (2015). Selection, selection, selection: the impact of return migration. *Journal of Population Economics*, pages 1–29.

Wahba, J. and Zenou, Y. (2012). Out of sight, out of mind: Migration, entrepreneurship and social capital. *Regional Science and Urban Economics*, 42(5):890 – 903.

Wakefield, M. A., Loken, B., and Hornik, R. C. (2010). Use of mass media campaigns to change health behaviour. *The Lancet*, 376(9748):1261 – 1271.

Wang, S. X. (2014). The effect of parental migration on the educational attainment of their left-behind children in rural China. *The BE Journal of Economic Analysis & Policy*, 14(3):1037–1080.

Willett, W. (1994). Diet and health: what should we eat? *Science*, 264(5158):532–537.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*.

Wu, D.-M. (1974). Alternative tests of independence between stochastic regressors and disturbances: Finite sample results. *Econometrica*, 42(3):529–546.

Yang, X., Derlega, V. J., and Luo, H. (2007). Migration, behaviour change and HIV/STD risks in China. *AIDS care*, 19(2):282–288.

Zhang, D., Meng, X., and Wang, D. (2010). The dynamic change in wage gap between

urban residents and rural migrants in Chinese cities.

Zhang, J. (2006). Wuqianwan liushoucunfu feizhengchang shengcun diaocha (survey on the life of 50 million left-behind women. *Zhongguo Jingji Zhoukan (China Economic Weekly)*, 40:186–200.

Zhang, J. and Zhao, Z. (2015). Social-family network and self-employment: evidence from temporary rural–urban migrants in China. *IZA Journal of Labor & Development*, 4(1):1–21.

Zhang, S. (2012). Migration and childrens health: Evidence from rural China. *University of Houston, Department of Economics Working Paper*.

Zhao, Y. (1999). Leaving the countryside: rural-to-urban migration decisions in China. *American Economic Review*, pages 281–286.

Zhao, Y. (2002). Causes and consequences of return migration: Recent evidence from China. *Journal of Comparative Economics*, 30(2):376 – 394.

Zhao, Y. (2003). The role of migrant networks in labor migration: The case of China. *Contemporary Economic Policy*, 21(4):500–511.

Zhongdong, M. and Guowei, Z. (2009). Isolated or compensated: The impact of temporary migration of adult children on the wellbeing of the elderly in rural China. *Geographical review of Japan series B*, 81(1):47–59.

Zhu, N. (2002). The impacts of income gaps on migration decisions in China. *China Economic Review*, 13(2):213–230.

Zhu, Y., Wu, Z., Wang, M., Du, Y., and Cai, F. (2012). Do migrants really save more? understanding the impact of remittances on savings in rural China. *Journal of Development Studies*, 48(5):654–672.

Zimmerman FJ, S. S. (2014). The effects of food advertising and cognitive load on food choices. *BMC Public Health*, 14:342.