The Income Body Weight Gradients
in the Developing Economy of China

Darjusch Tafreschi

The Income Body Weight Gradients in the Developing Economy of China[1]

Darjusch Tafreschi

Author's address: Darjusch Tafreschi
SEW-HSG, University of St.Gallen
Varnbüelstrasse 14
CH-9000 St. Gallen
Phone   +41 71 224 24 74
Email   darjusch.tafreschi@unisg.ch

Abstract

Though existing theories predict the income gradient of individual body weight to change sign from positive to negative in process of economic development, empirical evidence is scarce. This paper adds to the literature on that topic by investigating the case of China using data from the China Health and Nutrition survey. Using a one-dimensional measure to characterize the level of economic development of a region, regression analyses indicate that more income is related to larger future growth of individuals' BMI in less developed areas whereas it lowers BMI growth in more developed areas. The switch is somewhat more pronounced for females. Finally, using concentration indices it is shown that overweight status is predominantly a problem of higher income ranks in less developed geographical areas and trickles down to lower income ranks throughout the course of economic development.

Keywords

JEL Classification

1 Introduction

Overweight induced by excess body fat poses a serious threat to individuals' health states as it significantly increases the probabilities of non-communicable diseases such as cardio-vascular disease, diabetes, hypertension, musculoskeletal disorders (especially osteoarthritis) and various cancers (endometrial, breast and colon). The rise of overweight in developed societies throughout the last decades and its potential effects on public health is one of the most intensively studied subjects in the recent health economics literature (see e.g. Cutler et al. (2003); Rosin (2008); Philipson and Posner (2008); Bleich et al. (2008)) - and the debate is still ongoing. Determinants identified are technological progress (and directly related more sedentary lifestyles) (Philipson and Posner (2003); Lakdawalla and Philipson (2002)), changes in food prices (Schroeter et al. (2008)), increased fast-food availability and cigarette taxation (Chou et al. (2004); Currie et al. (2010)), sugar sweetened beverages (Pereira (2006)) and bounded individual knowledge about the potential health consequences (Kan and Tsai (2004)) of overweight.

On the contrary, the rise of the weight levels in developing countries has long been neglected by economists. At present, the World Health Organization (WHO) projects that by 2015, approximately 2.3 billion adults will be overweight and more than 700 million will be obese (Caballero (2007)) - there is no doubt that a significant share is contributed by individuals from developing countries. A recent paper by Sahn (2009) strengthens this point by showing that there are few countries which are not affected by rising weight levels. Out of his sample of 30 developing countries, 17 show first order statistical dominance when their intra-country weight distribution is compared to past weight distributions. In addition, second order dominance is observed in 11 cases. While epidemiologists like Popkin (1994, 1998, 1999) have clearly spotted the ongoing "nutritional transition" in the developing world already some time ago, economic research (traditionally more concerned about the consequences of under-nutrition) investigating the phenomenon of growing weight levels in low-income countries is still relatively scarce. Some piecewise evidence on the phenomenon of increasing body-weight levels in developing countries has been delivered by Galal (2002); Fernald (2007); Doak et al. (2006); Case and Menendez (2009); Du et al. (2004). In short, the papers commonly identify higher levels of obesity for the majority of developing countries, especially affected are females and more wealthy household. Surprisingly, Sahn and Younger (2009) find a significant share of households in which underweight and overweight individuals coexist. Their analysis shows that between one half and two-thirds of BMI inequality is accounted for by within-household BMI.

A striking difference between obesity in developed and developing societies has first been

pointed out in a meta-study by Sobal and Stunkard (1989) - and has later been updated and reconfirmed by Monteiro et al. (2004b): While overweight is mainly a problem of individuals of lower socio-economic levels in developed countries (negative correlation between income level body-weight) (see Ball and Crawford (2005)), the opposite is true for developing countries where overweight is relatively more widespread among individuals belonging to higher income classes (positive correlation between income level and body-weight). Empirical evidence for the latter can be found for the country specific cases of Mexico (Fernald (2007)), Egypt (Asfaw (2007)) and China (Popkin (1999)). Philipson and Posner (2003) nicely modeled this inverted U-shape claiming that *in poor or early societies the more obese are relatively wealthier, but in wealthy, more modern societies the more obese are relatively poorer*. In fact, some studies (e.g, Monteiro et al. (2004a); Sarlio-Lahteenkorva et al. (2004)) support the overall existence of such a hump shaped income-body-weight relationship. The underlying idea is that starting from a situation where overweight is mainly a problem of higher income classes, with economic development it trickles down to lower income classes. Since an individual body-weight is determined by the relative ratio of caloric intake to caloric expenditure an increase (decrease) of this ratio, ceteris paribus, will lead to a weight increase (decrease).. Income influences both, caloric intake and caloric expenditure. With respect to the former Schroeter et al. (2008) argue that a negative income-body-weight relationship can be achieved when higher income is used to substitute a high-caloric diet by a low-caloric one, ceteris paribus. In this case, calories would be a quasi-inferior good from a certain income level onwards and higher income individuals would show lower (or negative) weight growth. Regarding the relationship between income and caloric expenditure one might imagine a latent demand for thinness which comes into play increasing voluntary activity (e.g. sports activities) once a certain income level has been achieved. Whether weight growth becomes negative for higher income classes at a certain point is unclear and ultimately an empirical question. With respect to this question Philipson and Posner (2003) argue that although the income-body-weight gradient switches sign at a certain stage of economic development, average weight will nevertheless be higher in more developed societies, because there is less on the job-exercise given the nature of occupations in such environments (i.e. less jobs in physically demanding fields like manufacturing and more jobs in services).

While theories and piecewise evidence support the hypothesis of an income-body-weight gradient that changes from a positive to a negative sign with economic development, no study has ever shown such a relationship empirically in one go. This paper attempts to fill this gap by providing empirical insights into the relationship between individuals' incomes and their weights, and changes thereof with respect to economic development. The analysis is carried out using the example of China. Contrary to a more naive framework where the gradient of

interest would be compared between countries at different stages of economic development, the within country approach taken in this paper has the advantage of being subject to a more equal institutional setting as it exploits within country differentials of development. Micro level data from a highly informative longitudinal survey is utilized, namely the China Health and Nutrition Survey (CHNS). The use of longitudinal information allows a more dynamic analysis rather than just a snapshot of a single moment in time. In particular, China provides an ideal case to investigate that relationship for several reasons. Following the economic opening of the country in 1979, China's economy grew at an unprecedented average yearly rate of about 10% for the last two decades, and that momentum has been held steady (see Holz (2008)). At the same time, this development was rather unequally distributed over the country. While coastal regions were the main beneficiaries of the uprise of the Chinese economy, provinces in inner China were dramatically left behind (Yang (1999)). Heilig (2006) documents in a global comparison that in terms of regional GDP per capita some Chinese provinces are comparable to industrialized countries while others are still at levels of low-income countries. This provides a set-up where within the same country there is significant within and between variation with respect to the development stages of its provinces.

Parallel to economic development, China has experienced tremendous increases of overweight levels - especially in urban areas (see Popkin et al. (2006)) . Wu (2006) and Levine (2008) provide evidence that the nutrition transition in China is currently progressing at fast pace with already more than 20% of the Chinese being overweight and another 3% being obese nowadays - translating into more than 200 million Chinese being affected by the epidemic. Given tremendous (but unequal) economic growth and significant increases in body weight patterns at the same time allows to investigate how these two phenomena are interlinked, and whether throughout this process -as theories predict- there is a shift of the disease towards lower income groups. No doubt, an analysis of the relationship between income and body-weight-levels at different stages of development is highly relevant to policy makers as it helps to understand the dynamics between income and body-weight and also allows to identify sub-groups of the population most at risk. Thus, for policy makers to be able to design possible countermeasures a deeper understanding of the dynamics is key.

In what follows, the paper will first give some information about the broad background of the topic discussing recent trends in weight levels, nutrition intakes and levels of physical activity. Moreover, a first descriptive investigation is done on how weight levels vary across wealth categories and different levels of development. After that a description of the data used, namely the CHNS, is given. Then I will describe the method facilitated to investigate the research question. Section four describes how community data from the survey is

used to conceptualize the concept of economic development. The resulting one dimensional measure aims to characterize individuals' living environments at the time when surveyed. The reliability of the measure is discussed in the second part of the same section. Section five presents a simple econometric set-up for the analysis at the level of individuals. The sample is divided according to different strata of economic development and BMI growth equations are estimated based on observable characteristics. Section six utilizes the concept of concentration indices to give further support to the trickling down hypothesis. Section seven concludes this paper.

2 Background

China was economically frail for well over the first half of the twentieth century and hit the rock bottom between 1957 and 1962 when - known as the *Great Leap Forward* - governmental policies to boost industrial growth resulted in the worth famine in the history of mankind.[1] During this period food production was clearly insufficient. Correspondingly, food consumption was generally low and while undernutrition was wide-spread, little evidence about increases of overweight is known. For most of that time there was limited access to progressive technologies and occupations (e.g. in agriculture or heavy industry) were mainly characterized by demanding physical activities leaving no space for excess body weight. While the economy remained at low levels throughout it was just in 1978 when broad structural reforms[2] opened the path for future economic development of the country. Subsequently, China's GDP grew on average 9.6% per year between 1978 and 2005 - and growth still persists on comparable levels to date (9.1% in 2009). Yet, developmental progress has been unequally distributed over the country and the drivers were mainly the coastal regions (see Holz (2008)).

While this economic success story is unprecedented there is consensus that it was accompanied by certain negative health effects on the population as documented by Van de Poel et al. (2009). As predicted by the theory of Philipson and Posner (2003) China's economic and technological progress also came along with significant increases in average body weight and is now a major concern to health policy makers. Alarmingly, Popkin et al. (2006) estimate the future health cost of the overweight epidemic (and direct consequences thereof) to reach 9% of China's GDP by 2025. A WHO report finds that China could loose $558 billion of national income till 2015 due to diabetes and heart disease - both illnesses are known

---

[1] Jisheng (2010) estimates this devastating famine to have claimed 36 million lives!

[2] One key element of the 1978 reforms was the liberalization of food production. Moreover, foreign investments were stimulated.

to be closely related to excess body weight. The main drivers of this increase are reduced physical activity of individuals due to less physically demanding occupations[3] and lifestyles (see Ng et al. (2009)) as well as a nutritional shifts (see Du et al. (2004)). Table 1 provides descriptive insights about this process by showing gender-specific mean trends of body size measures as well as of nutritional behaviours and caloric expenditures.[4] As common in the obesity literature, this paper uses the body mass index (BMI) measure (i.e. body-weight relative to squared height (kg/m^2)) to describe individual body shapes. By WHO standards an individual is considered as overweight or obese if his BMI exceeds 25 or 30 respectively.

Looking at broad averages the increase of body weights is inevitable. For both, females and males, average BMIs have risen continuously throughout the 15 years under consideration - for the latest years, mean BMIs were well beyond 23. While the averages are still about 3-4 points lower than those for developed countries (e.g. Germany or the U.S.) in absolute terms, BMI growth rates in China are already comparably higher (see Popkin (2008)). Correspondingly, the percentages of overweight and obese females and males have risen too. In 2006, close to one-fourth of females and males aged 18-60 were statistically classified as overweight. Comparing the figures to Sahn (2009) this is still relatively moderate in international comparison. However, given the large population of China, obesity rates of this order translate into more than 200 million adults being affected at present.

Yet, rises in BMI wrongly suggest daily calorie intake to have increased over the same time span. Surprisingly, the numbers contrarily reveal that mean caloric intake (kcal) during the same time has actually *decreased*. Roughly speaking, the data indicate that both sexes on average consume about 500 kcal less per day[5] which I interpret as a result of changing caloric requirements. However, the table also shows that, parallel to reductions in total caloric intake, the relative composition in individual's diets have clearly shifted away from proteins and carbohydrates towards foods rich in fat. This is equally true for females and males respectively. In fact, China has traditionally been regarded as a country with a relatively lean population that commonly featured a healthy diet rich in cereals and vegetables. This, however, appears to be changing rapidly. A shift away from the classic diet which is reflected by the data has also been found by Du et al. (2004). The numbers suggest that the country is undergoing a nutritional transition at high speed in which its traditional diet is continuously

---

[3]According to Holz (2008) China's share of laborers in agriculture declined from just above 70% to 45% between 1978 and 2005.

[4]Two remarks are in order. First, the means are calculated based on survey observations of CHNS (i.e. only 9 Chinese provinces are covered) and are, thus, not nationally representative. Second, some of the observed changes might be due to aging effects of the cohorts under consideration. Average age increases by approximately 1,5 years between the survey waves. However, the same calculations been made in a regression framework holding age constant yielding basically the same trends.

[5]This is in line with findings of Lakdawalla and Philipson (2002) for the U.S. who find constant (or even decreasing) levels of caloric intakes over a period of significant increases of BMI levels.

Table 1: Means of Selected Key Variables by Gender, 1991-2006

	Females 18-60						Males 18-60					
	1991	1993	1997	2000	2004	2006	1991	1993	1997	2000	2004	2006
Body Weight Measures												
BMI	21.97	22.05	22.49	22.97	23.16	23.13	21.52	21.68	22.16	22.70	23.10	23.23
Overweight (BMI>25)	0.13	0.14	0.18	0.21	0.23	0.23	0.09	0.10	0.14	0.20	0.23	0.24
Obese (BMI>30)	0.02	0.01	0.02	0.03	0.04	0.03	0.01	0.01	0.01	0.02	0.03	0.03
Caloric Composition												
kcal	2519.61	2441.35	2226.01	2223.46	2166.26	2073.00	2916.49	2813.93	2603.14	2587.94	2520.42	2457.63
fat (g)	61.31	62.93	64.58	71.73	70.39	66.84	68.45	69.82	72.25	80.05	78.55	75.62
protein (g)	72.58	71.11	64.90	64.58	63.65	63.53	83.82	81.94	75.81	74.93	73.71	74.48
carbohydrate (g)	416.39	391.72	344.16	321.97	314.73	302.24	471.77	444.32	400.00	372.03	359.90	350.91
Caloric Expenditure												
Employed	0.86	0.86	0.81	0.76	0.63	0.63	0.94	0.93	0.91	0.87	0.80	0.79
Light Activtiy	0.29	0.27	0.35	0.37	0.43	0.47	0.25	0.22	0.28	0.29	0.33	0.34
Moderate Activity	0.16	0.15	0.15	0.12	0.17	0.14	0.21	0.23	0.20	0.20	0.21	0.22
Heavy Activity	0.53	0.57	0.48	0.47	0.38	0.39	0.53	0.54	0.50	0.48	0.45	0.43
No Activity	0.01	0.01	0.02	0.04	0.02	0.01	0.01	0.01	0.02	0.03	0.02	0.01
1st Income Quintile												
BMI	21.51	21.76	22.07	22.67	22.96	23.01	21.06	21.33	21.79	22.16	22.45	22.77
% overweight	0.09	0.11	0.13	0.18	0.22	0.22	0.04	0.08	0.1	0.13	0.16	0.19
2nd Income Quintile												
BMI	21.63	21.85	22.06	22.7	23.07	23.17	21.06	21.35	21.68	22.45	22.82	22.79
% overweight	0.12	0.12	0.15	0.2	0.23	0.23	0.05	0.08	0.1	0.17	0.18	0.18
3rd Income Quintile												
BMI	22.05	21.94	22.5	22.95	23.19	23.32	21.46	21.68	22.07	22.31	22.71	23.02
% overweight	0.14	0.12	0.18	0.23	0.23	0.24	0.09	0.09	0.15	0.16	0.2	0.21
4th Income Quintile												
BMI	22.34	22.32	22.83	23.09	23.47	23.18	22.09	21.86	22.6	23.07	23.59	23.47
% overweight	0.15	0.16	0.21	0.23	0.26	0.23	0.13	0.11	0.17	0.23	0.27	0.26
5th Income Quintile												
BMI	22.34	22.4	22.97	23.36	23.17	23.01	21.97	22.17	22.66	23.5	23.93	24.06
% overweight	0.17	0.19	0.21	0.24	0.24	0.22	0.13	0.16	0.2	0.29	0.32	0.33
Non-Coastal Rural												
BMI	21.39	21.64	21.98	22.52	22.66	22.79	20.84	21.18	21.68	22.21	22.52	22.83
% overweight	0.09	0.1	0.13	0.16	0.18	0.19	0.04	0.06	0.1	0.14	0.17	0.19
Non-Coastal Urban												
BMI	21.97	22.11	22.68	23.04	23.08	22.95	21.52	21.81	22.26	22.58	23.03	23.06
% overweight	0.14	0.15	0.19	0.23	0.23	0.2	0.1	0.11	0.17	0.19	0.22	0.21
Coastal Rural												
BMI	22.5	22.35	23.15	23.42	23.83	23.93	22.13	22.18	22.79	23.26	23.66	23.85
% overweight	0.15	0.16	0.24	0.25	0.29	0.3	0.12	0.13	0.19	0.24	0.28	0.3
Coastal Urban												
BMI	22.61	22.69	23.64	23.65	23.73	23.58	22.45	22.55	23.56	23.87	24.15	24.21
% overweight	0.2	0.2	0.25	0.31	0.31	0.31	0.15	0.19	0.25	0.34	0.33	0.36

Source: Own Calculations.

giving way to more western diets characterized by animal foods and higher shares of edible oils.

Observing increasing body weights while actual caloric consumption dropped suggests that caloric expenditure (i.e. physical activity) must have decreased disproportionately.[6] In fact, this is what can be observed when examining indicators related to energy expenditure. Here, the percentage of employed individuals has steadily decreased over time. Furthermore, employment shares are consistently lower for females than for males. This is in parts due to longer spells in education.[7] Besides education an increasing share reports to seek work or -especially in case of females- do housework activities. Further insights can be gained by looking at indicators directly reflecting individual physical activity where three main categories are presented. Light activity includes jobs that are generally less physically demanding (i.e. sedentary jobs (e.g. office workers), teacher, laboratory technician, etc.). Moderate activity is defined to include occupations such as electrician, metal worker and salesman. Heavy activity, eventually, incorporates jobs which are physically demanding, e.g. farmers, steel workers, loaders, miners or stone-cutters. The time trends in these three categories are unambiguous: there is a marked shift away from heavy physical activities towards more light and sedentary ones. As with employment status, the shift is more pronounced for females where, by 2006, nearly half were classified to follow a light activity. In line with Popkin (1999) I interpret these changes as sectoral relocations brought about by the China's economic development, i.e. occupations are incrementally shifting from the agricultural and manufacturing sectors (both of which are characterized by relatively strenuous jobs) to the service sector (jobs here are often less physically demanding and sedentary). Hence, lower levels of physical activity appear to be the main driver of increasing weight levels. Following the arguments of Philipson and Posner (2003) this is most likely due to the fact that caloric expenditure has become relatively more expensive over time.

Investigating mean weight measures by income classes the table further shows that increases in BMIs were rather unequal. Apparently, back in 1991 there was a rather large disparity in overweight rates between lower and higher income quintiles. Average BMI and shares of overweight increased across income quintiles and both measures were consistently higher for females at that time - overweight rates were roughly twice and three times higher in the highest quintile than in the lowest for females and males respectively. As we have already seen overall weight levels rose in the following years, but for females the differences in average

---

[6]Scholars are in agreement that genetic changes in humans play -at most- a marginal role in explaining recent upward trends in body weights around the globe (see e.g. Popkin (1994)). Thus, the only reasonable explanation when observing increasing body weights while caloric consumption has decreased is a disproportionate decrease in caloric expenditure.

[7]While only 8% of the 18-26 year old claimed to be in education in 1991 the corresponding number more than doubled to 17% in 2006 (not shown in table).

BMI and overweight rates across income classes vanished. At the end of the observation period, in 2006, overweight rates for females have been homogeneously distributed throughout the income quintiles (with the middle quintiles even showing slightly higher levels). This resulted from two dynamics: weight levels rose at fast pace in lower income quintiles while they only grew slowly (or even stagnated and recently dropped) in the upper parts of the income distribution. Thus, while overall levels increased, there has been a shift of overweight concentration from higher income quintiles to lower ones for females. Naive extrapolation of that trend would suggest that in the future overweight levels will be higher for females in lower income quintiles. Yet, the situation is different for males. Here, too, all parts of the income distribution show considerable growth in average BMI and overweight over time. However, contrary to females, weight levels in males grew more heavily in higher income quintiles. By 2006 one-third of males at the end of the income distribution were classified as overweight - while in the lower quintiles this was only true for about 20%.

These major differences between females and males are consistent with what can be observed from the activity patterns. As perhaps in most societies, individuals at the bottom end of the income distribution in China are more likely to be involved in heavy activities. Since the movement out of heavy activities was more pronounced for females it is likely that they also experienced larger weight gains. For males, still a larger proportion is engaged in heavy activities preventing them from larger weight gains. However, why weight levels for males in higher income quintiles rose much faster than for females remains open.

The last set of rows in the table display mean BMI measures by geographical clusters which are assumed to reflect the varying levels of economic development. Four geographical categories are chosen, namely non-coastal rural, non-coastal urban, coastal rural and coastal rural areas.[8] Surely, this coarse classification does not fully capture the heterogeneity in terms of economic development within China. However, there is broad consensus that the main drivers of China's recent economic boom are urban areas, and especially those in coastal provinces (see Bramall (2009) for a comprehensive discussion). Hence, a comparison along these lines appears meaningful for the time being. Figure 8.2 (appendix) underlines existing differences between the geographical clusters in terms of per capita income. In addition to income levels, the areas show significant heterogeneity in terms of individual´s physical activity levels. Overall, between 1991-2006 the share of individuals with only light physical activity has increased steadily. However, it can be seen that especially individuals in urban areas are reducing their activity levels, a point that will be discussed in more detail later. As with different income quintiles, significant variations in mean weight levels are visible.

---

[8]The group of coastal provinces consist of Liaoning, Jiangsu and Shandong while non-coastal provinces are represented by Heilongjiang, Henan, Hubei, Hunan, Guangxi and Guizhou. See the data section for a corresponding map.

8

Throughout the time span under consideration coastal areas, and especially urban areas, show higher mean BMIs and, likewise, higher levels of overweight than their non-coastal counterparts. Again, this is true for males and females. By the year 2000 (and beyond), overweight prevalence in coastal urban areas was >30% for both genders. Inevitably, there has also been a pronounced growth in BMI levels in non-coastal provinces, but their absolute levels in 2006 were just as high as those of coastal areas in the beginning of the 1990s. Yet, investigating *changes* rather than absolute levels, growth rates have lately been most pronounced in rural areas.

3   Data

To empirically test the switching income gradient hypothesis of Philipson and Posner, this paper uses data from the China Health and Nutrition Survey (CHNS).[9] CHNS is an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. Unlike the naming of the survey may suggest it is *not* nationally representative, but only covers nine provinces: Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong[10] (see Figure 1). These provinces account for approximately 56% of the Chinese population. Within each of this provinces a draw of 24 communities were initially selected at random to participate in the survey providing a total number of over 200 communities that were grasped. When possible, these communities were constantly covered in subsequent CHNS rounds.

CHNS is of longitudinal nature and comprises questions at the individual, the household and the community level.. The same households were traced throughout the years of the survey and repeatedly interviewed when possible. So far, seven rounds of the CHNS are available to researchers namely 1989, 1991, 1993, 1997, 2000, 2004 and 2006 - yet the 1989 wave is not used in this study as then data was only collected for pre-schoolers and adults aged 20-45. Follow-up levels of the CHNS are high. Individuals are repeatedly followed and interviewed as long as they remain in their original community - but households that migrate to outside their community are not followed. Survey questions asked to individuals and their households are among others related to general demographics, time use, health indicators, educational outcomes and economic indicators such as labor force participation and various incomes. With respect to the latter, aggregate household income measures are

---

[9]For a broader discussion of the database and the sampling design, see Liu (2008). To access the official website, visit `http://www.cpc.unc.edu/projects/china`

[10]Note that Heilongjiang was first affiliated to the CHNS 1997. At the same time Liaoning did not participate in the 1997 wave, but returned for the consecutive waves.

Figure 1: CHNS Regions

created based on individually earned incomes (i.e. wages from employment, income from agricultural activities as well as from other types of self-employment less of investments in that business), unearned incomes (i.e. income derived from assets) and state-transfers (e.g. subsidy payments) respectively. This paper uses deflated[11] per capita household income instead of individual incomes as it is strongly believed that nutrition decisions are largely made at the household level. With respect to the outcome variables of interest, individuals' anthropometrics are exactly measured by trained health workers largely ruling out measurement error[12] and reporting bias respectively (see Cawley (2004); Chou et al. (2004)). Moreover, the fact that individuals are repeatedly measured allows to calculate changes in BMI over time. Furthermore, CHNS monitors individual and household food consumption in-depth allowing extensive insights into nutrition patterns - and inferences about energy intake (implying details on carbohydrates, fat and proteins).

At the communal level, officials are interviewed and a wide range of community specific information on issues like infrastructure, local labour markets, service availability, health facilities etc. is collected in each wave. Due to the panel structure these information allow to capture changes in community characteristics over time and, thus, to make judgments about developmental changes. Furthermore, a distinction between rural or urban communities has been made initially in the first wave of the survey where roughly two third of the communities where classified as rural. Yet, this variable has been held fixed throughout the later waves. As

---

[11]All incomes measures are inflated by the consumer price index using 2006 as the basis year.

[12]Remaining errors have been cleaned out during data processing.

this study aims to exploit not only between variation, but also within variation of communal characteristics, the rural-urban indicator does not suffice for this purpose.

The target population investigated in this paper consists of adults 18-60 years old at the time of the survey - thus children and elderly outside this age bracket are excluded from the sample. Moreover, from the outset I exclude observations with missing anthropometric data and BMI values smaller than 15 or larger than 40 (as those are extreme cases). Observations with negative values in real per capita household income were dropped from the sample.[13] Finally, cases were deleted for which there is no information on the communal level, after all leaving a unbalanced sample of 43,137 person-year observations (22,547 females and 20,580 males). Selected descriptive statistics for the sample of individuals are given in Table 2. To allow for gender heterogeneity in the relationship between income and body-weight all individual level analyses will be conducted for males and females separately.

4   Development Index

Construction of One-Dimensional Index of Economic-Development

This section describes how community information in the CHNS is used to construct a one dimensional index of economic development which is used in the later analysis. Although economists often approximate stages of development by simply looking at average income (per capita) I argue that for the purpose of this study a purely monetary measure of regional development is overly simplistic. Rating a region's relative standing solely based on the income of its inhabitants does not allow to fully capture the structural differences between regions that potentially affect the relationship of interest. Variations in terms of infrastructure, job market characteristics and available services could all potentially influence the way individuals at different positions of the income distribution behave with respect to their caloric intake and expenditure behaviours respectively. In the fashion of Van de Poel et al. (2009) and Jones-Smith and Popkin (2010) this paper will therefore define development in a broader manner embracing various community characteristics which are believed to reasonably depict a regions relative stage of development - among others also including average income. Advantageously, the CHNS provides a large battery of variables related to the surveyed communities that can be exploited for this purpose.

The variables used to assess a community's relative level of development are shown in

---

[13]This is done although negative incomes are reasonable under certain circumstances. Negative incomes might be observed if individuals report more expenses than income which can be legitimate if, for example if someone invests in a new small business that is not yet profitable at the time of the survey. I conducted all analyses without dropping these observation, the results remained very similar.

Table 2: Descriptive Statistics for CHNS Sample, Adults 18-60

Variable	Female, N=22,547		Male, N=20,580		Total, N=43,137	
	mean	sd	mean	sd	mean	sd
BMI	22.62	3.22	22.39	2.99	22.51	3.12
Overweight (BMI>25)	0.19	0.39	0.16	0.37	0.18	0.38
logIncome	8.02	0.98	8.04	0.98	8.03	0.98
Age	39.47	11.21	39.3	11.49	39.39	11.34
Activity Level						
Very Light	0.14	0.35	0.14	0.35	0.14	0.35
Light	0.22	0.41	0.14	0.35	0.18	0.39
Moderate	0.15	0.36	0.21	0.41	0.18	0.38
Heavy	0.47	0.5	0.49	0.5	0.48	0.5
Education						
No Education	0.27	0.44	0.11	0.31	0.19	0.39
Primary	0.22	0.42	0.22	0.42	0.22	0.42
Secondary	0.3	0.46	0.39	0.49	0.35	0.48
Higher	0.2	0.4	0.27	0.45	0.23	0.42
Marital Status						
Never Married	0.11	0.31	0.17	0.37	0.13	0.34
Married	0.85	0.35	0.81	0.39	0.83	0.37
Divorced/Seperated/Widoed	0.04	0.19	0.02	0.15	0.03	0.17
Year						
1991	0.17	0.38	0.17	0.38	0.17	0.38
1993	0.16	0.37	0.16	0.37	0.16	0.37
1997	0.16	0.37	0.17	0.37	0.16	0.37
2000	0.18	0.38	0.18	0.38	0.18	0.38
2004	0.16	0.37	0.16	0.37	0.16	0.37
2006	0.16	0.37	0.16	0.36	0.16	0.37
Geographical Cluster						
Rural Non-Coast	0.44	0.5	0.45	0.5	0.45	0.5
Urban Non-Coast	0.22	0.42	0.22	0.42	0.22	0.42
Rural Coast	0.23	0.42	0.22	0.41	0.23	0.42
Urban Coast	0.1	0.3	0.1	0.3	0.1	0.3
Share Non-Agricultural	58.96	34.89	58.14	35.04	58.57	34.96
Household Size	4.18	1.47	4.14	1.44	4.16	1.45
Dependency Ratio	0.26	0.2	0.24	0.21	0.25	0.21

Source: Own Calculations.

Table 3. Admittedly, the choice of the variables is to a certain extent driven by their availability throughout the community survey questionnaires. Potential candidate questions that were, however, not asked for at least one survey year have been excluded as their inclusion would have resulted in the loss of significant amounts of data. The inclusion of various other (partly related) measures has been carefully considered, but the selected ones have finally been chosen for mainly three reasons. First, the variables are believed to adequately capture the underlying development process as they cover various relevant areas. Second, while many other variables show a high degree of missing values, these variables are filled for most communities throughout all years of the survey. Third, carefully investigating within community variation of the variables, the selected candidates showed reasonable consistency over the years. [14]

The averages of the selected variables by survey year are itemized in Table 4. As can be seen, all the variables show the expected trends over time and their accuracy is supported by their consistent patterns. Simply judging from these means, apparently, the communities have become more developed over the years: infrastructure has improved, services have become more widely available, the labour markets are increasingly moving away from traditional agricultural jobs to ones characterized by less physical activity, regions have become more economically integrated and, finally, real average incomes have clearly risen between 1991 and 2006.

Yet, the choice to approximate the stage of economic development by multiple dimensions disallows a unique mapping without further assumptions. In order to make well-defined inter-communal comparisons of economic development feasible a single index measure is required. Following Van de Poel et al. (2009)[15] factor analysis retaining the first factor is utilized using the pooled sample of communities to form a single development indicator based on the community variables.[16] The advantage of factor analysis in this setting is that it allows to create a measure which -assuming that the underlying latent process that drives the overtime changes of the selected community variables is economic development- is not merely a product of weights *subjectively* assigned to each of the variables. Though factor analysis is a well-accepted tool for the purpose of dimension reduction[17] its resulting factors do not have a

---

[14]Apparent inconsistencies have been cleaned in the data. For example, missing values of continuous variables have been imputed by taking the average of the previous and the consecutive survey waves where possible. Moreover, binary variables showing clear inconsistencies compared to previous and consecutive episodes have been adjusted.

[15]Notice that Van de Poel et al. (2009) do not investigate the phenomenon of economic development, but the degree of urbanity. While these two concept might to some extent be similar, the final list of variables used in the factor analysis of this paper differs from that of their paper.

[16]Only the first factor has an eigenvalue significantly greater than 1.

[17]See Sahn and Stifel (2000) for a more detailed discussion on the use of factor analysis for approximation of wealth indicators

Table 3: Community variables reflecting level of development

Category	Variable	Description
Community Infrastructure	Dirt Roads	Community has mainly dirt roads
	Stone roads	Community has mainly stone or gravel raods
	Paved roads	Community has mainly paved roads
	Bus	Community has bus stop (or long distance bus stop)
	Train station	Community has a train station
Available Services in Community	Telephone	Community has a convenient telephone service
	Post	Postal service available in community
	Newsletter	Community received provincial newspaper the day it is published
	Electric power cut	Average days per week community has electric power cut-off
Indicators of Occupational Environment	Farmland	Community has farmland
	Share non-agricultural	Share of non-agricultural workers in community
	Very light activity	Share of respondents in communtiy claiming very light physical activity
	Light activity	Share of respondents in communtiy claiming light physical activity
	Moderate activity	Share of respondents in communtiy claiming moderate physical activity
	Heavy activity	Share of respondents in communtiy claiming heavy physical activity
Economic Location	Open Trade Area	Community is near open trade area/city ($<$ 2hr by bus)?
Wealth	Income	Mean per capita household income (inflated)

Source: Own Calculations.

meaningful interpretation in absolute terms. Nevertheless, the approach allows to deduce an ordinal ranking. In our case the resulting indicator (i.e. the first factor) allows both, cross sectional and over time comparisons of communities covered by the CHNS with respect to their stage of economic development. Using only community observations for which none of the selected variables are missing in a certain wave leaves us with 185 communities in 1991, 179 in 1993, 186 in 1997, 212 in 2000, 214 in 2004 and 213 in 2006 -altogether leaving 1189 community-year observations- that are used in the construction of the development index. Accordingly, individual level data can be enriched with the level of community development.

Table 4: Means of Community Variables by Year

Variable	Mean by Wave					
	1991	**1993**	**1997**	**2000**	**2004**	**2006**
Dirt Roads	0.25	0.23	0.19	0.13	0.05	0.03
Stone Roads	0.14	0.17	0.17	0.14	0.15	0.13
Paved Roads	0.47	0.55	0.59	0.68	0.77	0.84
Bus stop	0.54	0.59	0.67	0.77	0.78	0.83
Train station	0.14	0.18	0.24	0.21	0.23	0.28
Telephone	0.63	0.72	0.84	0.93	0.96	1
Post	0.85	0.9	0.95	0.96	0.96	0.97
Newsletter	0.32	0.4	0.47	0.51	0.63	0.76
Electric power cut	1.11	0.67	0.37	0.15	0.12	0.2
Farmland	0.65	0.63	0.58	0.54	0.53	0.5
Share non-agricultural	0.54	0.58	0.59	0.61	0.66	0.68
Very Light Activity	0.11	0.1	0.16	0.17	0.2	0.25
Light Avtivity	0.14	0.13	0.16	0.16	0.2	0.19
Moderate Activity	0.3	0.31	0.29	0.28	0.25	0.22
Heavy Activity	0.35	0.36	0.33	0.33	0.29	0.3
Near Open Trade Area	0.25	0.33	0.35	0.41	0.44	0.4
Average Income	2561.5	2918.83	3672.64	4774.03	6402.98	7392.47

Source: Own Calculations.

The result of the factor analysis reveals that about 57% of the common variance among the input variables is explained by the retained first factor. The time constant factor loadings, i.e. the degree to which each of the community variables correlates with our factor of interest, are given in Table 5. The loadings are bounded between -1 and 1 and their absolute values indicate the corresponding correlation with the underlying factor. All loadings show the expected sign (in line with theoretical arguments) fortifying the assumption that the underlying first factor indeed reflects the relative level of economic development of communities. Clearly, enhancements in infrastructure and transport, service availability, economic integration and higher average incomes correlate positively, while labour markets characterized by more heavy physical occupations and a high concentration in the agricultural sector univocally reflect relatively lower degrees of development.

Table 5: Factor Loadings of Community Variables

Variable	Factor Loading
Dirt Roads	-0.50
Stone Roads	-0.31
Paved Roads	0.69
Bus stop	0.36
Train station	0.23
Telephone	0.45
Post	0.25
Newsletter	0.50
Electric power cut	-0.31
Farmland	-0.74
Share non-agricultural	0.80
Very Light Activity	0.78
Light Avtivity	0.73
Moderate Activity	0.11
Heavy Activity	-0.90
Open Trade Area	0.32
Income	0.54

So far, the index variable resulting from the factor analysis is originally continuous in nature. However, for the development index to be used as a mean to stratify the data with regard to the later analysis, a discrete measure is required. Therefore, an additional discrete variable is created that is directly based on the distribution quintiles of the originally continuous measure. This discrete variable classifies the development stage of a community as being in the categories from one to three with the highest value representing the highest level of development. The decision to choose exactly three classes of development stages is directed by two arguments. On the one hand there should be enough variation (i.e. there should be not too few strata) to allow the capturing of effect heterogenieties in the main analysis. On the other hand, a large number of strata reduces the number of observations in each of the strata and, thus, reduces the precision of our result[18].

Assessment of Development Index

For the subsequent analysis to be credible it is essential to assure trustworthiness of the development index. Following DeVellis (2003), particularly two properties of the index should be ensured. First, it should reflect the over-time development that China has experienced for the last two decades, i.e. the index should on average show higher values for later periods than for earlier ones. Second, as there is certainly heterogeneity between communities and

---

[18]All analyses have been conducted in addition on four and five development strata with the qualitative results remaining largely the same.

regions, the index should be able to discriminate accordingly.

Plotting the average logarithmic incomes of communities against their value of the development index by waves provides Figure 2. The vertical lines are drawn at the mean value of the overall development index, i.e. at zero, while the horizontal lines are drawn at the mean value of the logarithm of average income (i.e. at 8.26). Linear fit-lines are added. The graph consistently shows a positive relationship for all waves- partly determined by the inclusion of income in the construction of the index - but at the same time revealing that communities with similar values in their development index have significant variation with respect to their income levels. In line with the magnitude of the factor loading this supports the assumption that apart from income also non-monetary variables should be considered in the assessment of communities development levels. Moreover, it can be seen that in time the scatter-plot is monotonously moving from the lower-left quadrant to the upper right one which is an indication of increasing levels of aggregate development over time. This point is further underscored by Figure 3 which draws the naive kernel density estimates of the distribution of the development index for each of the waves. Two things are noteworthy. First, with time the estimated distribution of the development index is clearly shifting rightward, once again supporting the hypothesis of overall developmental progress. Second, the densities of all years are bimodal in nature. This is probably due to persisting rural urban disparities. In fact, when following the definition of rural and urban areas as they are initially coded in the CHNS community surveys, the development index discriminates surprisingly well between them. In the group of communities that are assigned to be rural the average of the development index does not exceed the zero (mean) value in any of the waves while for the urban sub-group the average never falls below it.

Still, significant variation of the development index even within the sub-groups of rural and urban communities indicates that the index is able to capture nuances beyond a crude rural urban definition. To further highlight how the index captures variation in the stage of development between communities, Tables 6 and 7 show unweighted averages of the constructed development index for each of the covered provinces and for geographical clusters (as defined in section 2) by year.

Again in line with the findings from above, time variation suggests that on the aggregate all provinces exhibited a positive trend in their development paths. With the exception of Shandong in 1993 the development index universally increases over time.[19] Investigating differences across provinces it can be seen that the eastern coastal regions (especially Liaon-

---

[19]Note that in Shandong province in 1993 only 16 communities were covered by the CHNS, i.e. 7-8 communities fewer than throughout all other years and compared to other provinces respectively. Moreover, the missing communities in 1993 were mainly classified as urban in the previous wave. Hence, it is likely that the missing of these communities causes the development index to temporarily drop in 1993.

Figure 2: Logarithmic Income vs. Development Index by Wave
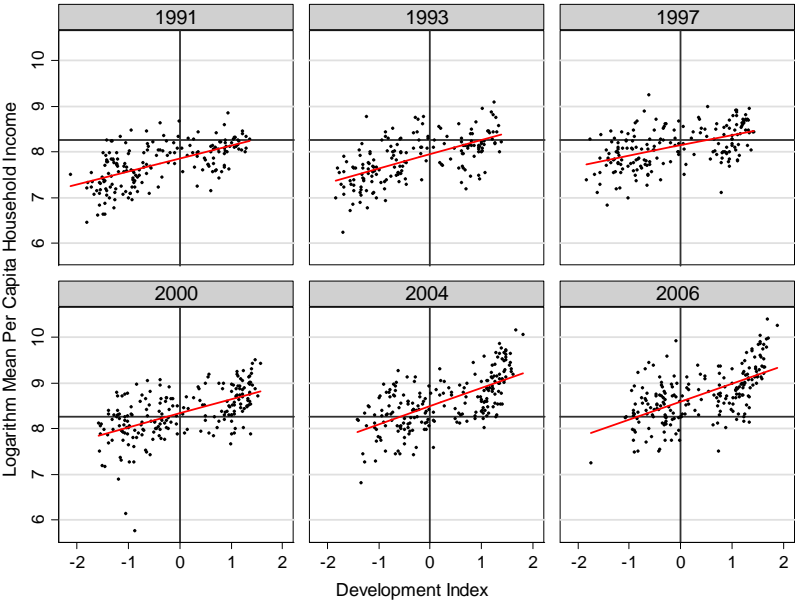
Note: Own calculations. The line represents the unconditional linear fit between the two measures.

Figure 3: Distribution of Development Index by Survey Year

Note: Own calculations.

18
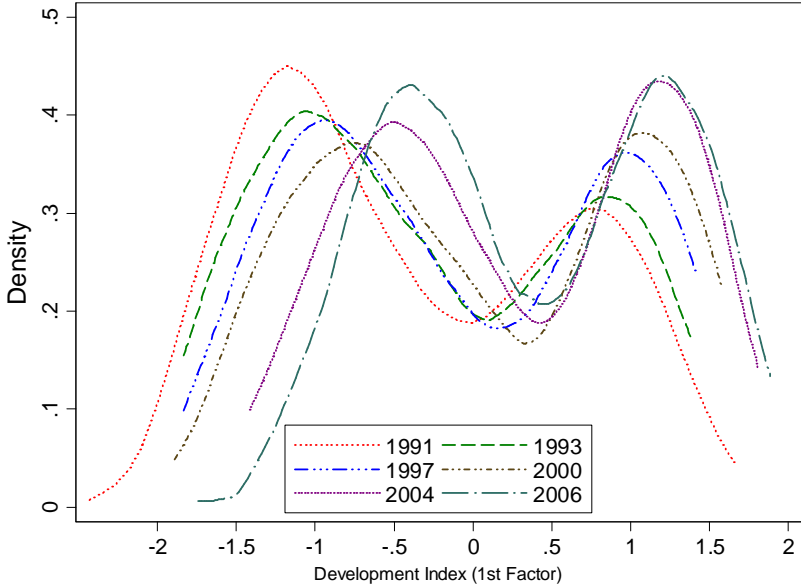
Table 6: Average Development by Province and Year

			Year			
Province	**1991**	**1993**	**1997**	**2000**	**2004**	**2006**
Liaoning	0.01	0.03		0.22	0.43	0.64
Heilongjiang			-0.37	-0.22	-0.01	0.14
Jiangsu	-0.41	-0.29	0.11	0.12	0.41	0.44
Shandong	0.05	-0.12	0.26	0.36	0.53	0.63
Henan	-0.79	-0.59	-0.3	-0.17	0.12	0.24
Hubei	-0.47	-0.34	-0.19	-0.15	0.13	0.21
Hunan	-0.5	-0.2	0.03	0.12	0.48	0.54
Guangxi	-0.43	-0.26	-0.13	0.11	0.25	0.31
Guizhou	-0.6	-0.4	-0.3	-0.14	0.07	0.21

Table 7: Development Index by Geographical Clusters and Time

			Year			
Area	**1991**	**1993**	**1997**	**2000**	**2004**	**2006**
Non-Coastal Rural	-0.84	-0.68	-0.48	-0.32	-0.07	0.06
Non-Coastal Urban	-0.10	0.24	0.42	0.54	0.72	0.84
Coastal Rural	-0.63	-0.56	-0.20	-0.22	0.10	0.20
Coastal Urban	0.76	0.84	0.91	1.07	1.18	1.24

ing and Shandong) appear to be comparatively more developed throughout the entire time span under consideration. This point is further supported by looking at mean levels of the development index by geographical clusters. As discussed earlier, urban areas exhibit higher levels of development throughout the time span under consideration. Moreover, especially coastal regions are consistently at the forefront. At the same time regions in inner China consistently show lower levels in their values for all waves considered which supports the ability of the index to capture between variations.

Finally, Table 8 presents the communities' transition probabilities with respect to the development quintiles between two consecutive waves labeled as t and $t + 1$.[20] For this, a modified version of the discretized development index variable is used. E.g. the entry in the second column of the first row gives the transition probability that anyone community was classified to belong to the lowest development quartile in one wave, but jumped to the next higher (i.e. the second lowest) quartile in the subsequent wave. Accordingly, the entry in the third column of the third row represents the probability that anyone community was classified to belong to the third quartile and remained in that also in the subsequent wave. Assuming that the majority of communities experienced positive economic development throughout

---

[20]There are exceptions where communities are left out for one wave. In that case consecutive waves refers to the two closest waves in which the community is observed.

the time under considerations one should expect to find that pattern to be reflected by the data. Indeed, the figures confirm that the majority of communities experienced constant or positive development stages, i.e. probabilities of remaining in a development category or jumping into a higher category are higher in absolute terms. Moreover, the higher the category in the initial period t, the more likely it is that a community stay in the same period in the consecutive wave. At the same time it can be seen that there is a significant number of communities that - given their initial development stage- show a jump into a higher category in the consecutive period. Only few cases ($< 5\%$) experienced a drop in the classification.

Table 8: Development Quintile Transition Probabilities in %

		DQ t+1				
		1	**2**	**3**	**4**	**5**
DQ t	**1**	60.09	32.19	7.73	0	0
	2	8.9	54.97	35.08	1.05	0
	3	1.16	11.63	65.12	19.19	2.91
	4	0	0	2.09	69.11	28.8
	5	0	0	0	11.59	88.41

In sum, the constructed index appears to reasonably reflect communities' *relative* stage of development. It shows some desirable properties in that it reenacts the overall positive development path that most Chinese areas have undergone during the last 20 years. Moreover, it accentuates between community differentials in development consistently. In what follows, the created index is used to characterize individuals economic environment and to proxy the average level of development within regional sub-clusters. These sub-clusters are at varying aggregation levels. The roughest aggregation level is at province-wave level, the finest aggregation at community-wave level. At rougher levels of aggregation the mean of the development index of the comprised communities is assumed to reflect the level of economic development in that cluster.

5 Individual Level Analysis

Set-Up

This section investigates the determinants of individuals' *BMI growth* with special focus on the income gradient and whether it is any different when comparing less developed areas with their more developed counterparts. The level of development is empirically approximated

by the categorization based on community variables as described in the previous section.[21] Focusing on changes of BMIs (instead of levels) implies that only adult individuals can be included in the analysis which are captured at least twice by the CHNS. Reducing the sample in that way leaves 13,980 female and 12,408 male (26,388 total) observations.[22] For the remaining observations I calculate the future annual growth rate in their BMIs utilizing the two nearest measurements. Using OLS regression, individuals' future BMI growth between period t_0 and period t is modeled as follows:

$$BMI_{growth,i} = \beta_0 + \beta_1 BMI_{t_0,i} + \beta_2 logInc_{t_0,i} + \beta_3 PA_{t_0,i} + \beta_4 X_{t_0.i} + \beta_5 wave_{t_0} + \beta_6 geo_i + \varepsilon_i$$

In this specification the geometric mean growth of height adjusted body weight ($BMI_{i,growth}$[23]) is a linear function of its present level (BMI_{t_0}), the log of deflated per capita household income ($logInc_{i,t_0}$), the present level of physical activity (PA_{i,t_0}), additional covariates (such as age at the time of the survey and its square, education, marital status, household characteristics) as well as time and geography dummies. As usual, ε_i is the remaining error term. Other studies that analyze BMI as a function of income usually focus on the present level of body weight and *not* on its growth. Hence, the individual data (oftentimes of cross-sectional nature) used usually measure BMI and income simultaneously. In those studies the relationship between income and BMI is either estimated as the conditional correlation between the two (e.g. Jolliffe (2010), Lu and Goldman (2010) or Chou et al. (2004)) or is identified as a causal effect using instrumental variable approaches (Akee et al. (2010) or Cawley et al. (2010)). It is easy to see that judgments about the relationship of interest are difficult to make in such settings.

On the one hand, even though all confounding variables are controlled for, potential reverse causality between income and BMI makes coefficient estimates questionable.[24] Thus, identification strategies that estimate the effect of $logInc_{t_0}$ on BMI_{t_0} are likely to fail even when controlling for observable characteristics of individuals yields conditional independence. On the other hand, although overcoming biases due to (1) omitted variables in the estimated equation and (2) reverse causality, approaches that investigate long-lasting exogenous variations in incomes to establish causal parameters have the major shortcoming that they are

---

[21]Preliminary analyses (not shown here, but available from the author upon request) that stratified subsamples according to geographical clusters (non-coastal rural, non-coastal urban, coastal rural and coastal rural areas) instead of the development index yield qualitatively similar results.

[22]For the survey wave of 2006 only individuals' BMI can be used to calculate the growth rates.

[23]Formally, the geometric mean growth rate of BMI is defined as $BMI_{growth,i} = ((\frac{BMI_{t,i}}{BMI_{t_0,i}})^{\frac{1}{t-t_0}} - 1)$.

[24]Shimokawa (2008) actually shows that (at least at higher levels of the body-weight distribution) there is a negative effect of weight on income.

not able to make general statements, because the effect is only identified for specific sub-groups in the population.[25] While these local effects are interesting in their own right, policy makers are usually interested in a broader understanding of the problem. Moreover, in the present context it is difficult to think of a meaningful instrumental variable. Income per se has no direct effect on BMI, but only through resulting changes of caloric intake and caloric expenditure. Hence, for exogenous changes in income to have a measurable effect on BMI the effect needs to be long-lasting. However, for the exclusion restriction to hold, any potential IV candidate would need to affect BMI only through the income channel. Yet, it is just natural to assume that exogenous shifts in income will also initiate changes in the level of physical activity (substitution effects). While it is possible to control for the present level of physical activity, it is per se not possible to control for the changes thereof as induced by variations in the instrument thus making any IV strategy likely to fail.

Analyzing *future BMI growth* (instead of its level) as a function of income is a somewhat different approach - yet very informative. By using the panel structure provided by the CHNS we can learn how the present level of household income -holding other factors constant- relates to the magnitude of future changes in individuals' body weights. There are several advantages of using future BMI growth instead of its present level. First, from an econometric point of view potential problems due to the reverse relationship between income and BMI are ruled out by construction in this set up as it is difficult to think of a person's future growth in body weight to have any influence on his income level today. Second, since the outcome represents a percentage change, the β_2 coefficients can be interpreted as elasticities. Third, policy makers are mostly interested in the dynamics of the obesity epidemic and less in snapshots. Modeling future changes in BMI allows to identify sub-populations most at risk of future increases of BMI and also whether the dynamics are any different across females and males and different levels of economic development, respectively.

It is important to note that changes in body weight over time can be considered as a process of convergence, because there is a natural upper bound to an individual's body weight. Moreover, it is likely that factors such as income will have different effects on $BMI_{i,gwt}$ depending on its current level. Hence, it is essential to control for BMI_{t_0}[26] when modeling its future growth. Convergence effects are expected to yield negative values for β_1 independent of the level of economic development. Likewise, BMI is an increasing and concave function of age which, in turn, is likely to affect the income level of individuals. After controlling for the level of physical activity and other potential confounders, β_2, is assumed to indicate

---

[25]It is well-known that instrumental variables only allow to identify the effect of interest for the sub-population of compliers, i.e. the part of the population that reacts to the instrument.

[26]This approach is similar to that of growth models in macro-economics where controlling for present levels of GDP is common practice in GDP growth equations.

whether (in expectation) a higher level of household income is related to larger or smaller future growth in BMI respectively. The analysis is conducted for each gender separately using geography and time fixed effects. Since the model is estimated on a pooled sample the standard errors are adjusted for clustering at the individual level.

Results

Table 9 shows the pooled OLS estimates for the determinants of BMI_{gwt} by gender and level of economic development under the inclusion of geography and year dummies. The mean values of female and male annual growth rates for adults 18-60 in the data are 0.48 and 0.57 respectively. The constant terms reveal that for both genders BMI_{gwt} has been most pronounced in the most developed areas.

The coefficients of main interest are to be found in the first row and clearly support the hypothesis of a rotating income gradient. When looking at the different sub-samples, in the least developed areas a higher level of (log) income is related to a significantly larger future growth in individual BMI. This is true for males and females likewise, and the coefficients are similar in size. After controlling for the other covariates in the model a 10% increase in income is, on average, related to a 0.015 p.p. (0.013 p.p. for males) increase in BMI_{gwt}. Moreover, the income gradient of looses its significance when looking at the middle development category (though the point estimates are still positive). It should be noted that here the coefficients are approximately half the size of their correspondents in the least developed categories. The overall picture changes drastically for the economically most developed areas where the income gradients switch from positive to negative. Here, for females a 10% increase in income is related to a 0.017 p.p. *decrease* in BMI_{gwt}. Stated differently, in more developed areas the body weights of females located at higher levels of the income distribution (ceteris paribus) grow significantly *slower* than of those at lower levels - a pattern which is qualitatively similar to what can be in industrialized societies nowadays. Although the income gradient is also negative for males in the most developed areas it is not significantly different from zero.

At this point it is also worth to investigate the other coefficients of the model. As expected, there are strong convergence effects of BMI_{gwt} as reflected by the significantly negative coefficients across all sub-samples- thus, higher levels of BMI_{t_0} are followed by significantly lower future growth rates[27] across all stages of economic development . Yet, while the convergence effect for females is strongest in the most developed areas the contrary is true for males where the least developed areas reflect the largest degree of convergence of BMI_{gwt} w.r.t. its initial

---

[27] Interestingly, the pattern of BMI convergence also hold when absolute annual changes of BMI (not shown here) are regressed on the same set of covariates. Hence, convergence effects found are not only due to the way the dependent variable BMI_{gwt} is constructed, namely as being a function of BMI_{t_0}.

Table 9: OLS Results by Gender and Development Level, Adults 18-60

	FEMALES			MALES		
	Dev.Cat=1	Dev.Cat=2	Dev.Cat=3	Dev.Cat=1	Dev.Cat=2	Dev.Cat=3
Dep. Variable (in %)	BMI_{gwt}	BMI_{gwt}	BMI_{gwt}	BMI_{gwt}	BMI_{gwt}	BMI_{gwt}
$logInc_{t_0}$	0.1458***	0.0751	-0.1720**	0.1329***	0.0689	-0.0839
	(0.051)	(0.062)	(0.073)	(0.049)	(0.061)	(0.074)
BMI_{t_0}	-0.3302***	-0.3473***	-0.3990***	-0.4110***	-0.3018***	-0.3192***
	(0.021)	(0.026)	(0.023)	(0.031)	(0.022)	(0.021)
Age	0.1142***	0.1883***	0.1750***	-0.0093	-0.0579	-0.0205
	(0.039)	(0.043)	(0.053)	(0.035)	(0.040)	(0.054)
Age^2	-0.0014***	-0.0023***	-0.0017***	-0.0000	0.0004	0.0002
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
Not Working	reference	reference	reference	reference	reference	reference
Working	1.0848	0.3292	-0.1216	0.6246	0.6872	0.3238
	(0.681)	(0.330)	(0.218)	(0.817)	(0.464)	(0.270)
Very Light PA	reference	reference	reference	reference	reference	reference
Light PA	0.8175	0.6987**	-0.0995	0.1690	0.0246	0.3167
	(0.717)	(0.303)	(0.208)	(0.829)	(0.518)	(0.352)
Moderate PA	-0.1320	0.4563	0.3562	0.2709	0.3472	0.4672
	(0.732)	(0.353)	(0.333)	(0.908)	(0.507)	(0.378)
Heavy PA	0.4813	0.5367*	-0.6793	0.4226	0.3471	0.1861
	(0.663)	(0.294)	(0.760)	(0.808)	(0.504)	(0.722)
Never Married	reference	reference	reference	reference	reference	reference
Married	-0.1165	-0.1467	-0.3370	0.2186	0.4695**	0.1868
	(0.329)	(0.315)	(0.313)	(0.170)	(0.211)	(0.273)
Divorced/Widow	0.1450	-0.4114	-0.6780	-0.4685	0.2805	0.5630
	(0.454)	(0.422)	(0.430)	(0.321)	(0.401)	(0.402)
No Education	reference	reference	reference	reference	reference	reference
Primary	-0.0179	0.0562	0.1980	0.0828	-0.1420	0.2883
	(0.099)	(0.134)	(0.210)	(0.179)	(0.224)	(0.325)
Secondary	-0.0090	-0.0145	-0.0243	0.1363	-0.2530	0.4091
	(0.127)	(0.164)	(0.198)	(0.191)	(0.228)	(0.296)
Higher	-0.0930	-0.1232	-0.3383	0.1328	-0.3259	0.2460
	(0.182)	(0.198)	(0.217)	(0.269)	(0.259)	(0.301)
Constant	3.5272***	3.3334***	6.9035***	7.5023***	7.6474***	8.6822***
	(0.995)	(0.969)	(1.132)	(1.062)	(1.007)	(1.263)
Geography Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,903	4,330	3,747	5,314	3,823	3,271
R-squared	0.091	0.109	0.146	0.116	0.093	0.118

Note: Robust standard errors in parentheses. Standard errors account for ID clusters.

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

Table does not report coefficients and s.e. for the following variables: Interaction Terms of Working Status and Physical Activity Level, Education of Household Head, Household Size, Dependency Ratio, Share of Males in Housheould

level. Furthermore, the model hints at a non-linear relationship between Age and BMI_{gwt}. For females, the growth rates increase up to the age of 40 (and become negative at age 80) in the least and middle developed areas while reaching the maximum approximately at the age of 50 in the most developed areas. For males, interestingly, both Age coefficients are insignificant.

Astonishingly, most of the coefficients that capture the effect of caloric expenditure due to physical activity appear to be insignificant. This might be due to the fact that some variation in physical activity levels is already captured by the development index. A priory one would assume individuals with low levels of physical exhaustion (i.e. non-employed and with very light and light physical activity) to have larger values of BMI_{gwt}. This, however, is not reflected by the estimates.[28] If anything, the only two significant coefficients are those indicating that females in the middle development category with light and heavy physical activity of females exhibit *larger* values of BMI_{gwt} over the subsequent period than those with very light activity levels (the reference category). Although not in line with the a priory expectation this might still be reasonable. Economic development is a dynamic process which is accompanied (among other things) by a continuous move of the labour force out of physically demanding occupations (often agriculture and heavy manufacturing) into less demanding jobs (service sector). Or, stated differently, with economic development there is a trend that individuals are continuously lowering their caloric expenditure, ceteris paribus, due to the changing structure of the labour market (Lakdawalla et al. (2005)). As seen, many individuals covered by the CHNS also face such structural changes over the observed period. Accordingly, larger values of BMI_{gwt} among individuals with higher levels of physical activity -as found by the estimates- might simply be due to the fact that these individuals are likely to reduce their physical activity levels in the future to a greater extent than those who are in the very low activity category already and, therefore, gain more weight ceteris paribus.

In contrast to other studies, the level of education plays virtually no role in the process of body weight accumulation after controlling for the other variables. The same is basically true for the education of the household head. This finding is not in line with the prediction of theoretical models which consider the decision to be overweight as an outcome of limited knowledge about potential health risks. Finally, the relationship between marital status and BMI_{gwt} is also rather weak. However, while being married is related to larger future weight gains for males (though only significantly in the middle category) the reverse is true for females.

These findings, however, should not be confused. As seen in the descriptive section,

---

[28]The coefficients for the interaction terms of employment and activity level are not reported here, yet they are all insignificant

the overall level of average BMI is still a lot higher in more developed areas. Yet, when extrapolated, the changing gradients indicate the hump shape proposed by existing theories.

6 Concentration Indices Analyses

A positive (negative) income body-weight relationship can be understood as a higher (lower) mean level of body-weight at upper ranks of the income distribution. Thus, a non-zero gradient implies a unequal distribution of body-weight w.r.t. income levels. Especially in health economics the concentration index (Kakwani (1980)) has been extensively used to quantify and compare socio-economic status related inequalities in health variables. I follow this practice. The concentration index (CI) is closely related to the concentration curve[29] and is defined as $C = \frac{2}{\mu} cov(I, H)$, where μ represents the mean level of the health variable looked at in the distribution of interest, I the income measure and H the health measure. This estimator is also referred to as the convenient covariance formula and can easily be applied to individual (ungrouped) data such as the CHNS.[30] By definition, the concentration index ranges from -1 to 1 while zero indicates a completely equal spread of the health variable across the income distribution, $C < 0$ ($C > 0$) indicates that the health measure (i.e. higher body weight) is relatively more concentrated among the poorer (richer) ranks, i.e. larger absolute values indicate a more unequal distribution of the health variable. Applied separately at regional clusters in the data, ceteris paribus, CI can directly be used to compare the degree and the direction of health inequalities between these different sub-populations (i.e. cross-sectional and over time). Exactly for that reason, CI is, though simple, especially well-suited for the present analysis. Assuming that there is a rotating income gradient, this would imply that overweight is shifting relatively stronger towards relatively poorer individuals. In case of such trickling down of overweight one should observe a negative relationship between the level of economic development of a region and its CI. For the computation of the concentration index per capita household income deflated to 2006 prices is used as the ranking variable I.[31] Regarding H, the binary information on an individuals overweight status (=1 if overweight) is used. Wagstaff (2005) has shown that when the health variable of interest is

---

[29]The concentration curve plots the ranked (poorest to richest) cumulative percentage of the income variable on the x-axis against the cumulative percentage of the health variable on the y-axis.

[30]In Kakwani et al. (1997) the authors provide the regression equivalent (convenient regression) to estimate concentration indices based on individual data. This approach allows to control for confounding demographic factors (e.g. age) when computing CIs and, thus, to extract *avoidable inequalities* which are not due to demographic factors. However, as found in the analysis, controlling for potential demographic confounders hardly alters the results in the present context. Hence, all CI estimates are based on the convenient covariance formula without further adjustments.

[31]Note that the concentration index does not depend on the absolute variation in the ranking variable but only on its relative ranking.

binary, the bounds of the CI are contracted.[32] It is adviced to normalize the CI by dividing it through $1 - \mu$ in order to allow direct comparison of the CI across different sub-population. I follow that practice. All CIs are calculated for males and females separately.

Two counteracting effects are at play when defining the aggregation level at which CIs are computed. At rougher levels of aggregation the number of individual observations in the sub-cells are larger which leads to more precise estimates of the CI for the sub-populations. However, rougher aggregation levels also reduce the number of the resulting data points that can be used for comparisons. On the contrary, computing CI at finer aggregation levels will lead to more data points, but at the cost of precision. Fortunately, the CHNS structure allows to conduct the analysis at different degrees of aggregation. As with the development index, single communities at distinct points in time can be identified from the data. This is the finest level of aggregation at which it is reasonable to compute CI. Yet, the number of observations at this level is (very) low - in fact too low to form a basis for a meaningful computation of CI at that level[33]. For this reason, the sub-samples for the construction of CIs are based on a rougher level of aggregation. First, CIs of overweight are computed at the aggregation level of wave and gender specific sub-samples formed at the province level. Moreover, within each province the sample is further divided based on rural/urban status. This approach delivers 204 sub-samples, in each of these CIs are computed. Moreover, for each of these sub-samples the average development level is computed based on the development index at community levels. The relationship between these two measures is analyzed in a regression framework to shed more insights on the relationship between incomes and body-weight throughout the process of development. A simple scatter plot of CIs against the development levels (supported by linear fit lines) reveal some first evidence of a changing income gradient at the aggregate level.

The graph shows the negative correlation between a regions average level of development and its CI. It should be noted that the linear fit line for females appears to be at a lower level compared to the line for males. At a development index value 0.5 it crosses the critical zero-threshold. In other words, the CI indicates that the most developed areas covered by the CHNS are at a stage where overweight is already now a predominant problem of the lower

---

[32]The interval of the concentration index is shrunken to the lower and upper bounds $\mu - 1$ and $1 - \mu$ respectively. Thus, the larger the mean value of the health variable the smaller the bounds.

[33]When restricting the sample to adults 18-60 years old and dividing it further by gender, some community wave clusters comprise less than 20 observations. Preliminary analysis shows that CI standard errors, if computed at community wave levels, become too large to still be useful. Moreover, at this aggregation levels about 15% (346 of 2412) of the sub-clusters exhibit mean overweight levels of zero ($\mu = 0$) which makes computation of CI in these clusters infeasible.

Figure 4: CI vs Development Level
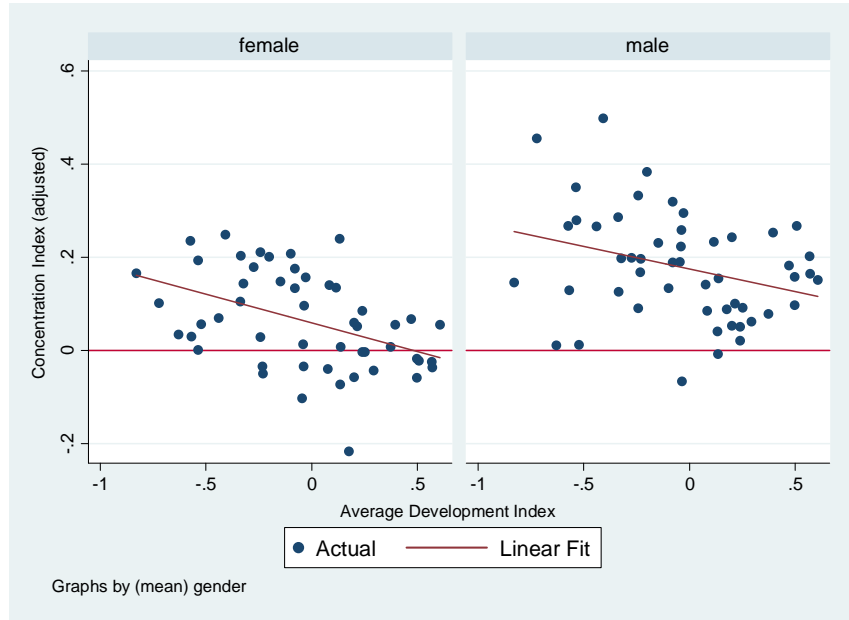
Note: Own calculations.

income classes. Clearly, this result is consistent with the finding from the previous section, namely that at later stages of economic development there is a stronger negative effect of income on weight gain for females.

However, the correlation might be contaminated by general time effects. To target this problem, I simply regress CIs on indicators of economic development as well as on dummies indicating the survey year. I further include a gender dummy that indicates if a CIs is computed for a male sub-population to see if there is a difference between genders. Table 10 provides results of regression analyses which explain the computed CIs as a function of economic development explicitly controlling for time and province specific effects. The first column reports the estimated coefficients of an regression of overweight CIs on a constant term and year dummies. It can be seen that throughout the survey waves the level of CIs has continuously decreased, for the final wave being on average 0.12 units lower than for the first one. Similar to Section 2, the naive geographic location is first used to approximate the level of development in the second column. The time dummies hardly change- however, the coefficients for the geography dummies show that in urban areas the average CI of overweight is found to be 0.09 units smaller than in rural non-coastal areas. The result hold when additionally controlling for gender. Among the sub-population of males, on average, the CIs are on average 0.09 points higher than for females. I ultimately interpret this as evidence that the income BMI gradient for males only switches at a later stage of economic development than for females (compare to Monteiro et al. (2004b)). The last two columns

replace the representation of development through geographic location by the continuous development index. In column four the coefficient is negative and significant (though only at the 10% level). When further adding province dummies to the model the negative effect of the development index becomes somewhat stronger and is now significant at the 5% level. Hence, all specifications indicate that there has been a trickling down of overweight to the lower ranks of the income distribution and the stage of development plays a crucial role in that shifting.

Table 10: OLS Results CI Overweight

VARIABLES	(1) CI_ow	(2) CI_ow	(3) CI_ow	(4) CI_ow	(5) CI_ow
Rural Non-Coast		reference	reference		
Urban Non-Coast		-0.0300	-0.0300		
		(0.0255)	(0.0240)		
Rural Coast		-0.0131	-0.0131		
		(0.0313)	(0.0295)		
Urban Coast		-0.0934***	-0.0934***		
		(0.0313)	(0.0295)		
Development Index				-0.0330*	-0.0370**
				(0.0177)	(0.0180)
Male			0.0996***	0.0992***	0.0992***
			(0.0196)	(0.0199)	(0.0185)
1991	reference	reference	reference	reference	reference
1993	-0.0542	-0.0542	-0.0542	-0.0490	-0.0484
	(0.0378)	(0.0372)	(0.0350)	(0.0356)	(0.0332)
1997	-0.0611	-0.0658*	-0.0658*	-0.0511	-0.0421
	(0.0378)	(0.0373)	(0.0351)	(0.0359)	(0.0342)
2000	-0.0616*	-0.0632*	-0.0632*	-0.0464	-0.0366
	(0.0367)	(0.0361)	(0.0340)	(0.0354)	(0.0337)
2004	-0.0925**	-0.0941***	-0.0941***	-0.0700*	-0.0594*
	(0.0367)	(0.0361)	(0.0340)	(0.0365)	(0.0349)
2006	-0.123***	-0.124***	-0.124***	-0.0963**	-0.0851**
	(0.0367)	(0.0361)	(0.0340)	(0.0373)	(0.0358)
Constant	0.156***	0.186***	0.136***	0.0971***	0.0868**
	(0.0267)	(0.0304)	(0.0303)	(0.0274)	(0.0374)
Province Dummies	No	No	No	No	Yes
Observations	204	204	204	204	204
R-squared	0.060	0.102	0.208	0.179	0.314

Standard errors in parentheses - *** $p<0.01$, ** $p<0.05$, * $p<0.1$
CIs computed at Province/Rural-Urban/Wave/Gender level.
CIs computed for adult populations and adjusted for demographic characteristics.

7   Concluding Remarks

This paper has shown that China is witnessing a drastic increase in BMI levels and, at the same time, a pronounced shift of overweight and obesity towards individuals at lower ranks of the income distribution. Descriptive statistics revealed that this comes despite reduced mean caloric intake and is therefore mainly driven by overcompensating reductions in physical activity. As modeled by Philipson and Posner (2003) I find that the income body-weight gradient rotates throughout the course of economic development whereas the latter is approximated using community related information of the CHNS. According to the results the relatively less developed areas covered by the CHNS are reflecting the typical "developing country patter": body-weights are growing relatively stronger at the upper ranks of the income distribution and overweight is so far predominantly a problem of relatively richer individuals in these areas. Yet, when investigating more developed areas I rather find an "industrialized country scenario" where a higher level of income is related to lower (or negative) growth of an individual's BMI ultimately effectuating the epidemic to trickle down to lower ranks of the distribution. Hence, the income-BMI gradient is not linear but hump-shaped and differs with the level of economic development. Moreover, for females the gradient appears to rotate at an earlier stage of development than for men.

References

AKEE, R., E. SIMEONOVA, W. COPELAND, A. ANGOLD, AND J. COSTELLO (2010): "Does More Money Make You Fat? The Effects of Quasi-Experimental Income Transfers on Adolescent and Young Adult Obesity," *IZA Discussion Paper No. 5135.*

ASFAW, A. (2007): "Do Government Food Price Policies Affect the Prevalence of Obesity? Empirical Evidence from Egypt," *World Development*, 35, 687–701.

BALL, K. AND D. CRAWFORD (2005): "Socioeconomic status and weight change in adults: a review," *Social Science & Medicine*, 60, 1987–2010.

BLEICH, S., D. CUTLER, C. MURRAY, AND A. ADAMS (2008): "Why is the developed world obese?" *Annual Review of Public Health*, 29, 273–295.

BRAMALL, C. (2009): *Chinese Economic Development*, Taylor & Francis.

CABALLERO, B. (2007): "The global epidemic of obesity: An overview," *Epidemiologic Reviews*, 29, 1–5.

CASE, A. AND A. MENENDEZ (2009): "Sex differences in obesity rates in poor countries: Evidence from South Africa," *Economics & Human Biology*, 7, 271–282.

CAWLEY, J. (2004): "The impact of obesity on wages," *Journal of Human Resources*, 39, 451–474.

CAWLEY, J., J. MORAN, AND K. SIMON (2010): "The impact of income on the weight of elderly Americans," *Health Economics*, 19, 979–993.

CHOU, S., M. GROSSMAN, AND H. SAFFER (2004): "An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System," *Journal of Health Economics*, 23, 565–587.

CURRIE, J., S. DELLAVIGNA, E. MORETTI, AND V. PATHANIA (2010): "The effect of fast food restaurants on obesity and weight gain," *American Economic Journal: Economic Policy*, 2, 32–63.

CUTLER, D., E. GLAESER, AND J. SHAPIRO (2003): "Why have Americans become more obese?" *Journal of Economic Perspectives*, 93–118.

DEVELLIS, R. (2003): *Scale Development: Theory and Applications*, Sage Publications, Inc.

DOAK, C., L. ADAIR, M. BENTLEY, Z. FENGYING, AND B. POPKIN (2006): "The Underweight/Overweight Household: An Exploration of Household Sociodemographic and Dietary Factors in China," *Public Health Nutrition*, 5, 215–221.

DU, S., T. MROZ, F. ZHAI, AND B. POPKIN (2004): "Rapid income growth adversely affects diet quality in China–particularly for the poor!" *Social Science & Medicine*, 59, 1505–1515.

FERNALD, L. (2007): "Socio-economic status and body mass index in low-income Mexican adults," *Social Science & Medicine*, 64, 2030–2042.

GALAL, O. M. (2002): "The nutrition transition in Egypt: Obesity, undernutrition and the food consumption context," *Public Health Nutrition*, 5, 141–148.

HEILIG, G. (2006): "Many Chinas? The Economic Diversity of China's Provinces," *Population and Development Review*, 32, 147–161.

HOLZ, C. (2008): "China's economic growth 1978-2025: What we know today about China's economic growth tomorrow," *World Development*, 36, 1665–1691.

JISHENG, Y. (2010): "The fatal politics of the PRC's great leap famine: The preface to tombstone," *Journal of Contemporary China*, 19, 755–776.

JOLLIFFE, D. (2010): "Overweight and poor? On the relationship between income and the body mass index," *IZA Discussion Paper No. 5366*.

JONES-SMITH, J. AND B. POPKIN (2010): "Understanding community context and adult health changes in China: Development of an urbanicity scale," *Social Science & Medicine*, 71, 1436–1446.

KAKWANI, N. (1980): *Income inequality and poverty: Methods of estimation and policy applications*, Oxford University Press Oxford.

KAKWANI, N., A. WAGSTAFF, AND E. VAN DOORSLAER (1997): "Socioeconomic inequalities in health: Measurement, computation, and statistical inference," *Journal of Econometrics*, 77, 87–103.

KAN, K. AND W. TSAI (2004): "Obesity and risk knowledge," *Journal of Health Economics*, 23, 907–934.

LAKDAWALLA, D. AND T. PHILIPSON (2002): "The growth of obesity and technological change: A theoretical and empirical examination," *NBER Working Paper No. 8946*.

LAKDAWALLA, D., T. PHILIPSON, AND J. BHATTACHARYA (2005): "Welfare-enhancing technological change and the growth of obesity," *American Economic Review*, 95, 253–257.

LEVINE, J. (2008): "Obesity in China: Causes and solutions," *Chinese Medical Journal*, 121, 1043–1050.

LIU, H. (2008): "The China health and nutrition survey: an important database for poverty and inequality research," *Journal of Economic Inequality*, 6, 373–376.

LU, Y. AND D. GOLDMAN (2010): "The Effects of Relative Food Prices on Obesity – Evidence from China: 1991-2006," *NBER Working Paper No. 15720*.

MONTEIRO, C., W. CONDE, B. LU, AND B. POPKIN (2004a): "Obesity and inequities in health in the developing world," *International Journal of Obesity*, 28, 1181–1186.

MONTEIRO, C., E. MOURA, W. CONDE, AND B. POPKIN (2004b): "Socioeconomic status and obesity in adult populations of developing countries: A review," *Bulletin of the World Health Organization*, 82, 940–946.

NG, S., E. NORTON, AND B. POPKIN (2009): "Why have physical activity levels declined among Chinese adults? Findings from the 1991-2006 China health and nutrition surveys," *Social Science & Medicine*, 68, 1305–1314.

PEREIRA, M. (2006): "The possible role of sugar-sweetened beverages in obesity etiology: A review of the evidence," *International Journal of Obesity*, 30, 28–36.

PHILIPSON, T. AND R. POSNER (2003): "The long-run growth in obesity as a function of technological change," *Perspectives in Biology and Medicine*, 46, 87.

——— (2008): "Is the obesity epidemic a public health problem? A decade of research on the economics of obesity," *NBER Working Paper No. 14010*.

POPKIN, B. (1994): "The nutrition transition in low-income countries: An emerging crisis," *Nutrition Reviews*, 52, 285–298.

——— (1998): "The nutrition transition and its health implications in lower-income countries," *Public Health Nutrition*, 1, 5–21.

——— (1999): "Urbanization, lifestyle changes and the nutrition transition," *World Development*, 27, 1905–1916.

——— (2008): "Will China's nutrition transition overwhelm its health care system and slow economic growth?" *Health Affairs*, 27, 1064–1076.

POPKIN, B., S. KIM, E. RUSEV, S. DU, AND C. ZIZZA (2006): "Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases," *Obesity Reviews*, 7, 271–293.

ROSIN, O. (2008): "The economic causes of obesity: A survey," *Journal of Economic Surveys*, 22, 617–647.

SAHN, D. (2009): "Weights on the rise: Where and for whom?" *Journal of Economic Inequality*, 7, 351–370.

SAHN, D. AND D. STIFEL (2000): "Poverty comparisons over time and across countries in Africa," *World Development*, 28, 2123–2155.

SAHN, D. AND S. YOUNGER (2009): "Measuring intra-household inequality: Explorations using the body mass index," *Health Economics*, 18, 13–36.

SARLIO-LAHTEENKORVA, S., K. SILVENTOINEN, AND E. LAHELMA (2004): "Relative weight and income at different levels of socioeconomic status," *American Journal of Public Health*, 94, 468.

SCHROETER, C., J. LUSK, AND W. TYNER (2008): "Determining the impact of food price and income changes on body weight," *Journal of Health Economics*, 27, 45–68.

SHIMOKAWA, S. (2008): "The labour market impact of body weight in China: A semiparametric analysis," *Applied Economics*, 40, 949–968.

SOBAL, J. AND A. J. STUNKARD (1989): "Socioeconomic status and obesity: A review of the literature." *Psychological Bulletin*, 105, 260–275.

VAN DE POEL, E., O. O'DONNELL, AND E. VAN DOORSLAER (2009): "Urbanization and the spread of diseases of affluence in China," *Economics & Human Biology*, 7, 200–216.

WAGSTAFF, A. (2005): "The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality," *Health Economics*, 14, 429–432.

WU, Y. (2006): "Overweight and obesity in China," *British Medical Journal*, 333, 362–363.

YANG, D. (1999): "Urban-biased policies and rising income inequality in China," *American Economic Review*, 89, 306–310.

Appendix

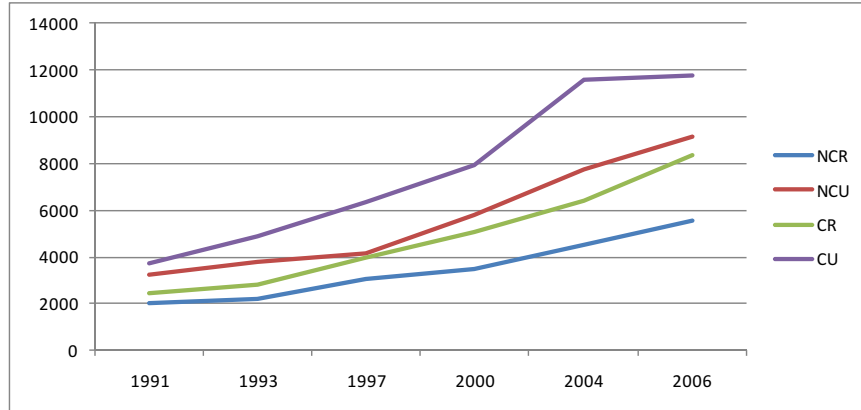Figure 5: CHNS Trends: Per Capita Income over Time for Geographical Clusters

Figure 6: CHNS Trends: Changes in Activity Levels over Time