HEDG Working Paper 08/25

Urbanization and the spread of diseases of affluence in China

Ellen Van der Poel

Owen O'Donnell

Eddy van Doorslaer

October 2008

ISSN 1751-1976

http://www.york.ac.uk/res/herc/research/hedg/wp.htm

Urbanization and the spread of diseases of affluence in China

Ellen Van de Poel[1*], Owen O'Donnell[2] and Eddy van Doorslaer[3,4]

[1] Erasmus School of Economics, Erasmus University Rotterdam
[2] University of Macedonia
[3] Erasmus School of Economics, Erasmus University Rotterdam
[4] Tinbergen Institute, The Netherlands

Abstract
A new methodology is used to quantify, track and explain the distribution of obesity and hypertension across areas differentiated by their degree of urbanicity. We construct an index of urbanicity from longitudinal data on community characteristics from the *China Health and Nutrition Survey* and compute a rank-based measure of inequality in disease risk factors by degree of urbanicity. Prevalence rates almost doubled over the period 1991-2004 and the risk factors became less concentrated in more urbanized areas. Decomposition analysis shows that urbanicity-related inequalities are mostly attributable to differences in community level characteristics and to disparities in incomes and in the physical and farming activity of individuals.

Keywords: China, urbanization, health inequalities, obesity, hypertension, decomposition

Acknowledgements
The data for this paper were supplied by the *China Health and Nutrition Survey*, funded by NIH (R01-HD30880, DK056350, and R01-HD38700) and the Carolina Population Center and the Chinese Centers for Disease Control. The authors are grateful to the Institute for Housing and Urban Development Studies (IHS) for funding the project on "Urbanization, Health and Health Inequality", from which this paper derives. Also many thanks to Alex Fu, Gordon Liu, Lars Osberg and Kuan Xu.

[*] Corresponding author :
Burg. Oudlaan 50, 3000 DR Rotterdam, The Netherlands
Tel: +31 10 4081477, Fax: +31 10 4089141, E-mail: vandepoel@few.eur.nl

Introduction

China will complete in just a few decades the urbanization process which took western developed countries three to four hundred years. China's urbanization rate was about 43% in 2005. Over the next 10-15 years, it is expected to rise to well over 50%, adding an additional 200 million mainly rural migrants to the current urban population of 560 million (Yusuf and Saich, 2008.) The consequences for population health are likely to be mixed. On the one hand, urban populations benefit from better access to health services, information and education, and have higher cash incomes and more economic opportunities. However, the rapid environmental, economic and social changes that follow urbanization increase the prevalence of major risk factors for chronic disease. In particular, urban areas in low and middle income countries are moving through a rapid nutritional transition towards Western-style diets, dominated by more processed foods and a higher fat content (Popkin, 2003; Popkin and Du, 2003). Increasing urbanization also leads to equally rapid shifts toward more sedentary occupations through the acquisition of new technology and transitions away from a mostly agricultural economy (Monda *et al*, 2007). In China, these transitions have contributed to stark increases in the prevalence of conditions such as obesity and hypertension, especially amongst male, urban and high income groups (Liu, 2004; Wang, Mi *et al*, 2005; Wang *et al*, 2007; Weng *et al* 2007). The emergence of non-communicable diseases as a major health threat in countries still coping with infectious diseases and childhood malnutrition threatens to overstretch already struggling health services. Forecasts estimate that China will lose $556 billion to the costs of heart disease, stroke, and diabetes in the period 2005–2015 (Wang *et al*, 2005).

Increasing urbanization and development is likely to drastically change the geographical distribution of these non-communicable diseases. This paper investigates how obesity and hypertension rates vary across areas at different stages of urbanization, and how and why this distribution is changing over time. In order to target public health interventions appropriately, it is important to establish whether these disease risk factors are spreading to less urban areas, or whether they are merely rising in the most urban ones. Knowledge of whether the geographic distribution of so-called diseases of affluence is changing because of changing population characteristics, behaviors or environmental factors is essential in identifying the type of interventions that are most likely to be effective in halting the spread of the diseases.

It is increasingly recognized that a simple urban-rural dichotomy is an oversimplification which cannot adequately distinguish the different living and health conditions experienced in areas at different stages of urbanization (McDade and Adair, 2001; Vlahov and Galea, 2002; Champion and Hugo, 2004; Dahly and Adair, 2004). Further, there is no universally agreed definition of "urban" and "rural", and in China the classification may have been distorted by the privileges to which non-agricultural residents were entitled (Kojima, 1995; Heilig, 1999). New criteria for the designation of new cities and towns have been introduced since 1983, resulting in changes of urban administrative areas. These have necessitated changes in the census definition of the urban population from time to time, causing much confusion in counting the number of urban Chinese (Wu 1994; Shen, 2006). Analysis of longitudinal survey data often presents a further problem in that the categorization of an area as urban or rural is fixed over survey waves. This is the case in the *China Nutrition and Health Survey* (CHNS) used here and so the dichotomous urban-rural variable does not capture the rapid urbanization of many designated rural areas that has occurred over the survey period. To overcome these problems, we construct an urbanicity index using community data from the CHNS[1]. This index gives a ranking of communities from low to high levels of urbanicity facilitating the use of rank-based inequality measures to examine the magnitude and trends of inequalities in obesity and hypertension across areas at various stages of urbanization. We also apply decomposition techniques to identify which determinants are driving these urbanicity related inequalities in obesity and hypertension, and their trends over time.

In the remainder we first present the CHNS data, conceptual framework and variable definitions used in the analysis. Thereafter we construct an urbanicity index and demonstrate its validity and reliability for the purpose of identifying communities at various stages of urbanization. We then illustrate how this index can be used to quantify urbanicity related inequalities in obesity and hypertension, and subsequently explain these inequalities and their trends by decomposing into the contributions of the determinants of the disease risk factors. The concluding section provides an interpretation of the implications of the study and its limitations.

Data

This study uses data from the CHNS, a large scale (and ongoing) longitudinal survey conducted in nine provinces of China in 1991, 1993, 1997, 2000 and 2004.[2] The provinces represented in the survey are Liaoning, Shandong, Jiangsu, Henan, Heilongjiang, Hubei, Hunan, Guangxi and Guizhou (shown in Figure 1).[3] Although the CHNS is not a nationally representative sample, the provinces covered vary substantially in terms of geography and economic development. The bulk of the Chinese population is located in the Eastern provinces, which are overrepresented in the CHNS, such that the selected provinces account for about 40% of the Chinese population (National Bureau of Statistics China, 2000). Since the pattern of city location in China has always been biased towards the East coast (Yan, 1990), our sample is selected from the relatively more urbanized regions of China, although it does not include Beijing and Shanghai, the most urbanized areas in China (and in the world). There is still substantial variation in urbanization rates across the CHNS provinces, ranging from 23% in Henan to 55% in Liaoning (year 2000 estimates by Shen (2006)). Urbanization rates also vary considerably within each province. When interpreting our results, one must bear in mind that we investigate inequalities by urbanicity in the CHNS data, which are not necessarily the same as those for China as a whole.

The CHNS collects information on a wide range of individual socioeconomic, health and nutritional characteristics, and – essential for this study – also detailed information on community characteristics. We examine the geographic distribution of disease risk factors by urbanicity in the first and last available wave of the CHNS and the change in the distribution that occurred over the intervening 13 year period during which China experienced tremendous economic development and urbanization. Our sample consists of 6484 adults (aged 16 years and older) in 1991 and 6197 in 2004, with about 45 individuals per community.[4] The health indicators of interest are obesity and hypertension, both defined as binary indicators. Individuals are considered hypertensive if their average of three systolic blood pressure measurements (at the time of survey) was equal to or greater than 140 mm Hg and/or average diastolic blood pressure was equal to or greater than 90 mm Hg and/or they were taking medication to lower blood pressure (Sixth report of the Joint National Committee on the Detection, Evaluation, and Treatment of Hypertension, 1997)[5].

Obesity is defined as a Body Mass index (BMI), based upon measured height and weight at the time of survey, above 25 kg/m^2.[6]

In selecting determinants of obesity and hypertension to be used to explain urbanicity related inequality in the prevalence of these conditions, we have followed the conceptual framework proposed by Northridge, Sclar and Biswas (2003) who distinguish between three groups of factors, on different levels, relevant to the association between urbanicity and population health. On the macro level, the framework includes the natural environment, macro social factors, and inequalities. Being macro-level conditions, these are relatively fixed for a country, or province, at one point in time and are generally only available at the country aggregate level. We include a set of provincial dummies in our regression models to capture province specific factors such as type of industry, climate, topography etc (Guizhou is the reference). Many studies have documented the higher prevalence rates of obesity and hypertension in the Northern regions of China, which seem to be associated with the higher salt intake and lower intake of fresh fruits and vegetables, as well as the colder weather in the north than in the south (He *et al*, 1995; Wu *et al*, 1995; Zhao *et al*, 2004; Weng *et al*, 2007).

These macro level factors, in turn, influence two domains of intermediate (community level) factors: the built environment and the social context. We consider these community level determinants as reflecting the urbanicity of a community. Table 1 shows a description of all community variables that we take into account, classified according to the framework of Northridge *et al* (2003). This list of community characteristics is generally very similar to that used by other studies (Liu *et al*, 2003; Monda *et al*, 2007). We have tried including more detailed information on health facilities, but as the coding of health facilities did not appear consistent across waves, we could only use availability of a health facility within the community. Further, we did not include population density of the community in the set of intermediate factors because of the high percentage of missing data.[7] However, since density is considered an important aspect of urbanicity, we have checked robustness of results to including density and reducing the sample size.

On the lower, more personal, level Northridge *et al* (2003) consider the proximate determinants, which consist of stressors, social integration and social support and health behaviors. Our set of proximate determinants (see Table 2) includes a set of age-gender dummies. Several studies have found the risk of obesity to increase with age, especially among women (Reynolds *et al*, 2007; Hou, 2008), whereas males are more at risk to be hypertensive (Wu *et al*, 1995; Hou, 2008). We also include marital status and a set of lifestyle variables: performance of heavy or very heavy physical activity, average daily fat intake (*logfat*)[8], smoking status (*smoke*) and whether the respondent consumes alcohol on a daily basis (*alcodaily*). Reynolds *et al* (2007) and Hou (2008) found that smoking, alcohol consumption and physical activity are the most important factors explaining regional differences in respectively obesity and hypertension in China. Also considerable evidence points to the importance of the transition from the Chinese low fat diet to the high fat diet of the West in explaining the increasing obesity and hypertension rates (Lukman, Dye and Blundell, 1998). While smoking has severe adverse health effects, and increases blood pressure, it is associated with lower levels of obesity (Ueshima, Zhang and Choudhury, 2000; Chou, Grossman and Saffer, 2004; Hou, 2008).

For socioeconomic status, we include a set of dummies for education level and (log) household income (*logincome*). The latter is calculated by summing all market earnings across the household and then adding the total value of all other non-market goods and services produced within that household (see Figure 1 in Liu *et al* (2008) for the breakdown of household income into its various components). Household income is then deflated using a year/province/urban-rural specific consumer price index that was developed for use with the CHNS[9], and divided by the (square root of the) total number of household members to obtain real household income per capita. The associations between socioeconomic status (income and education) and both obesity and hypertension have been found to differ according to the level of development. In higher income countries, obesity is more prevalent amongst the poor, whereas the reverse holds for low and middle income countries such as China (Wang, 2001; Kim, Symons and Popkin, 2004; Monteiro *et al*, 2004; Wang *et al*, 2006; Reynolds *et al*, 2007). Further, we control for individuals' economic activity status and occupation. More sedentary occupations such as office work are associated with an increased risk of obesity and hypertension as opposed to being involved in farming (Zhou *et al*, 2003; Monda *et al*, 2007).

Urbanicity index

Index construction

The concept of an urbanicity index was introduced in 1976 by Allen, and since then there have been a few attempts to develop an index from community level survey data. McDade and Adair (2001) use factor analysis on data from the Philippines, while Dahly and Adair (2007) subjectively assign weights to various community variables from the same data. For China, Liu *et al* (2003) create an index by equally weighting various community characteristics.

This paper applies factor analysis to a broad set of community level characteristics that can be expected to reflect a community's level of urbanicity. We prefer using factor analysis, as opposed to creating a score index that is based upon subjectively assigning weights to the various components, or putting equal weight on each component, as this guarantees that the weights are optimally chosen to maximize the explained variance in the underlying latent index.[10]

Although we will only investigate urbanicity related inequalities in 1991 and 2004, we do use all the waves in estimating the urbanicity index. Using as many observations as possible will increase the accuracy of the index. Table 3 shows summary statistics for the community variables across all waves from 1991 to 2004. Most variables show a trend that one would normally associate with increasing urbanicity: more communication and transportation possibilities, a move away from agriculture, more economic activity and more community services such as schooling and child care. There is a population drop from 1991 to 1997; however the difference between the waves is insignificant.

We retain the first factor, which we assume to reflect the urbanicity of a community. The urbanicity index explains 47% of the common variance in the community variables. Table 3 shows factor loadings of all these variables; all appear to have intuitive signs. Urbanicity is inversely associated with farming, bad road infrastructure, distance to markets and health facilities and electricity cut-offs. The negative association with availability of a primary school is counter-intuitive, but also very weak. By contrast, the urbanicity index correlates positively with more transport infrastructure, communication services, schools, child care facilities, and economic activity.

Assessing index reliability and validity

We now assess the validity and reliability of our urbanicity index (DeVellis, 2003; Dahly and Adair, 2007).

Internal consistency refers to the degree of interrelatedness of the items within the scale and is guaranteed by using factor analysis, which maximizes the amount of explained variance in the underlying latent index.

Temporal stability: If the urbanicity index is truly representative of the underlying latent variable, then it should consistently assess that latent construct at different points in time. As we can safely hypothesize that Chinese communities have become increasingly urban over the past decades, we expect our urbanicity index to increase over the waves of the CHNS as well. This is indeed confirmed: the average urbanicity index increases from -0.32 in 1991 to -0.13 in 1993, -0.07 in 1997, 0.09 in 2000 and 0.40 in 2004.

Criterion-related validity depends on the empirical association of the index with a ‘‘gold standard.’’ Although we have no gold standard of urbanicity (which is exactly the reason for constructing the index), we can gain some insight into criterion related validity by investigating how the index correlates with the subjective classification of each community as urban, suburb, town or rural by the CHNS interviewer in each wave. Figure 2 shows that our index does well in picking up different degrees of perceived urbanicity. Cities and towns have the highest average urbanization score, followed by suburban and rural areas. It is interesting that suburban areas clearly do not come second on the continuum from city to rural. This illustrates the usefulness of the index over a categorical approach. Figure 2 illustrates that all areas have become more urban during the respective time period, at a more or less equal pace. By 2004, there is no remaining difference in average urbanicity between towns and urban areas. This indicates that the infrastructure and services included in the urbanicity index have become equally likely to be present in towns as in urban areas. This suggests that our urbanicity index does not contain variables that can distinguish between highly urbanized areas (megacities) and towns with good infrastructure.

The upper panel of Figure 1 displays the average value of the urbanicity index for each CHNS province in 1991. The ranking of the provinces by these scores almost exactly corresponds to that estimated by Shen (2006, table 6), which provides further support for the criterion-related validity of the index.

Construct validity refers to the (theoretical) relationship between the urbanicity index and other variables. We investigate this by looking at the distribution of obesity, hypertension, fat intake and income across the distribution of the urbanicity index. As discussed before, it has been shown that all of these variables are more concentrated in urban areas. Table 5 confirms this, showing clear trends of adult obesity, hypertension, fat intake (grams), and income across quintiles of the urbanization index.

Added value of the urbanicity index

The disadvantage of relying on an urban–rural dichotomy is its inability to detect heterogeneity within urban and rural areas. For criterion related validity we want there to be a correlation between the urbanicity index and the CHNS categorization of communities as urban or rural, but, given difficulties referred to above in this categorization, the correlation should not be perfect and, for the index to provide considerable additional information, it should display substantial variation within each category of the dichotomy.[11] The distributions of the urbanicity index across urban and rural communities do overlap considerably, and there is substantial heterogeneity within categories (see Figure 3). The mean urbanicity index for urban communities (across all waves) was 0.52, with a range of [-1.7, 2.08], while the rural mean and range were -0.25 and [-2.50, 2.50], respectively. Overall, 17% of urban communities fell at or below the median urbanicity score, while 51% of rural communities lay above it.

In sum, our urbanicity index appears to be a reliable indicator of the degree of urbanicity of a community. It shows the expected time trends, correlates well with the categorical classification and the urban/rural dichotomy but it also reveals heterogeneity within the categories. And, unlike to the static urban-rural dichotomy, the urbanicity index enables us to track the – often dramatic – changes in the communities' environments over time. In the next sections, we use this index to quantify and decompose urbanicity-related inequalities in obesity and hypertension.

Urbanicity-related inequalities in obesity and hypertension

Measuring urbanicity related inequalities

Following Wagstaff, Paci and van Doorslaer (1991), the concentration index has been a popular measure of socioeconomic inequalities in health. Given that the concentration index is a rank-based measure of inequality, it can also be used to measure urbancity-related inequality with ranking provided by the urbanicity index. We use this approach to measure the degree to which obesity and hypertension are disproportionately concentrated among individuals located in more urbanized areas.

When applied to binary indicators such as obesity and hypertension, the bounds of the concentration index (CI) depend upon the mean of the indicator (Wagstaff, 2004). This impedes comparisons over time due to substantial differences in means across survey years. Average prevalence rates of obesity and hypertension have almost doubled over the period 1991-2004. In 1991, 13% of our sample was obese and 14% hypertensive, while by 2004 these percentages have risen to 26% and

23%, which illustrates the fast rising epidemic of chronic diseases in China. To avoid dependency of the CI bounds on the mean, we use an adjusted index suggested by Wagstaff (2004). This normalized concentration index is calculated as[12]:

$$CI_h = \frac{1}{1-H}\left[\frac{2}{H}\operatorname{cov}(h_i, R_i^u)\right] \quad (1)$$

where h_i is an indicator of whether the i-th individual suffers from obesity/hypertension (1/0), R_i^u is its respective fractional rank in the distribution of the urbanicity index and H represents average obesity/hypertension. It is important to note that given that our ranking variable—the urbanicity index—is defined at the community level, the concentration index could be calculated at that level using average community obesity/hypertension rates as the health outcome of interest. However, we prefer to keep the analysis at the individual level as this provides better estimates of the associations between obesity/hypertension and the proximate determinants, which will be used in the decomposition analysis below.

A positive concentration index means that the health outcome is more prevalent among people in more urban areas and vice versa.[13] For obesity, the normalized concentration index is 0.28 in 1991 and 0.11 in 2004, while for hypertension it went down from 0.21 to 0.14. So although in both years, obesity and hypertension are concentrated in more urban areas, the urbanicity-related inequalities have decreased over time. The prevalence of these conditions has increased tremendously over the period and they have become much less concentrated in the more urbanized areas. The use of the normalised CI ensures that the second result is not simply a consequence of this first. There are many potential explanations. It could be that the factors, such as a higher fat content diet and smoking that provoke these conditions are increasingly found in less urbanised areas. Alternatively, the spread of urbanisation itself could be responsible. Most communities have become more urbanized such that a given ranking in the distribution of the index represents a greater degree of urbanicity in 2004 than in 1991. Environmental conditions that generate problems of obesity and hypertension will be present at a lower ranking at the end of the period. With the aim of distinguishing between these and other explanations, we now turn to decomposition of urbanicity-related inequality.

Decomposing urbanicity related inequalities

Wagstaff, Watanabe and van Doorslaer (2003) have shown that the standard concentration index can be decomposed into inequalities in the health determinants. In a similar vain, a decomposition can be applied to the normalized index used here. The decomposition starts from the assumption that the health indicator of interest h_i can be written as a linear function of its determinants:

$$h_i = \alpha + \sum_{k=1}^{K} \beta_k x_{ik} + \varepsilon_i \qquad (2)$$

where we have assumed that h_i is associated with K observable covariates x_{ik}, and an idiosyncratic error ε_i. CI_h can be decomposed into inequality in its covariates as follows[14]:

$$CI_h = \frac{1}{1-H}\left[\sum_{k=1}^{K}\left(\frac{\beta_k X_k}{H}\right)CI_{x_k} + \frac{GC_\varepsilon}{H}\right] \qquad (3)$$

where X_k is the average of x_{ik}, CI_{x_k} are the (standard) concentration indices of the x_{ik} and GC_ε is the generalized concentration index of the residuals. The latter is calculated as $\frac{2}{n}\sum_{i=1}^{n}\varepsilon_i R_i^u$, and reflects the extent to which randomly distributed unobservables are contributing to urbanicity-related inequalities in h_i.[15]

A limitation of this decomposition method is that it only holds for linear models, while obesity and stunting are measured by binary variables, which are more appropriately modeled using non-linear estimators. While other approaches have been suggested in the literature to deal with this problem (O'Donnell, van Doorslaer and Wagstaff, 2006), we only show results using the linear probability model because it has the most intuitive interpretation, the alternative approaches have their own drawbacks and, in any case, because our results were found to be very robust to the method used.

Explaining urbanicity related inequalities in 1991

Table 6 shows regression and decomposition results for urbanicity-related inequalities in obesity and hypertension in 1991. The first column shows the (standard) concentration indices of the covariates and illustrates how these are distributed across the urbanicity index. Note that a positive concentration index means that the variable is disproportionately concentrated in more urbanized areas and a negative index indicates the opposite.

The signs of the concentration indices appear quite intuitive for most variables. We find that older people are more concentrated in urban areas, which might be related to the higher life expectancy in urban areas (Liu, Hsiao and Egglestone, 1999). Also, income and education are distributed pro-urban. Henan province is relatively rural, whereas Liaoning and Shandong are significantly more urbanized, as was also visible from Figure 1. Among the lifestyle variables, physical activity is distributed pro-rural and alcohol and fat intake are more concentrated in urban areas. Only farming is clearly more prevalent in the more rural areas, whereas all other occupations are concentrated in urban communities[16].

The second column of Table 6 shows the coefficients from the linear probability model with obesity as the dependent variable; the third column shows elasticities. Generally, the findings correspond to those found in the literature. Young people (below 30 years of age) have the lowest risk of being obese, females have a higher risk than men, and middle-aged women (between 45 and 65 years of age) are the most likely to be obese. Also, married individuals and those with higher incomes are more likely to suffer from obesity. High education on the other hand is protective against obesity. The province effects confirm the north-south gradient discussed before, with more obesity in Shandong, Henan and Liaoning than in Guizhou, and less obesity in Guangxi and Hunan. Among the life style variables we find that physical activity is protective against obesity while fat intake is significantly associated with a higher obesity risk. There is a negative association between smoking and obesity. All occupations are associated with lower obesity than the professional category; being a farmer is most protective. Not working and being a skilled worker has no significant effect. After controlling for all covariates, the urbanicity index is still highly positively correlated with obesity.

The fourth column of Table 6 show the results of the decomposition of urbanicity related inequalities in obesity (the fifth column shows grouped contributions of the same categorical variable). As shown in eq. (3), for a variable to contribute to urbanicity related inequalities, it has to be associated with obesity/hypertension, and its distribution needs to vary across the urbanicity index.

It is not a surprise that half of the inequality in obesity related to urbanicity is driven by the urbanicity index itself (50%), which captures most of the community characteristics of urban environments. This corresponds to the direct effect of urbanicity, as represented by community characteristics, on obesity that does not operate through the socio-demographic characteristics of the population, or its behavior. The next largest contribution comes from the occupation variables (29%), and particularly from being involved in farming (40%). That is, two-fifths of the inequality in the prevalence of obesity across more and less urbanized communities is explained by differences in engagement in farming given the strong negative association of this occupation with obesity. Higher incomes in more urbanized areas explain 11% of the inequality in obesity and differences in the fat content of the diet contribute 4%. About the same proportion of the inequality (5%) is attributable simply to the fact that the populations of more urbanized areas are older. Education lowers urbanicity related inequality (-11%) because it reduces the risk of obesity and urban populations are better educated. The province dummies jointly only contribute an insignificant 4%, suggesting that macro level factors are relatively unimportant in explaining differences in obesity by urbanicity. The contributions of the error term is negligible (-1%), indicating that our set of observed covariates does a good job in explaining urbanicity-related inequalities.

The regression results for hypertension shown in the sixth and seventh columns of Table 6 generally confirm the patterns observed for obesity. Hence, we only discuss those that differ. Hypertension increases with age and is more prevalent among males. Fat intake is not significantly associated with hypertension, and smoking is negatively associated. While cigarette smoking is a well-known risk factor for hypertension, other studies for China have also reported counter-intuitive results (He *et al*, 1995; Hou, 2008), which may derive from the strong positive correlation between smoking and socioeconomics that are not sufficiently captured by the education and income variables (Pan, 2004). Occupation is not significantly associated with hypertension, although the unemployed face a higher risk.

The eighth and ninth columns of Table 6 present the decomposition results of urbanicity-related inequalities in hypertension, which appear to be related to several factors. Differences in physical activity contribute most (29%) to inequality. More rural populations engage more in physical activity, which helps reduce the risk of hypertension. The contribution of the demographics is almost equally large (28%) and results from the fact that hypertension increases with age and older people are more concentrated in urban areas. The direct effect of the urbanicity index accounts for about a fifth (19%) of the inequality. Income differences explain 13% of the concentration of hypertension in more urbanized communities. The provincial dummies make a larger contribution (13%) than is the case for obesity, mostly deriving from Shandong and Liaoning being both relatively urbanized and suffering from higher hypertension rates. As was the case for obesity, the education advantage in urbanized areas helps suppress their greater risk of hypertension, reducing the inequality by 11%. Again, the contribution of the error term is quite small and insignificant (3%).

Urbanicity related inequalities in 2004

Table 7 shows regression and decomposition results for obesity and hypertension in 2004. There have been shifts in the distribution of demographics, with younger people being more concentrated in urban areas in 2004. This might be a consequence of age-selective rural-urban migration, with younger people moving away from rural areas (Wu, 1994). Further, there have been changes in the relative urbanicity of the provinces. Relative to the other provinces, Hunan and Guangxi have become more urbanized while the opposite happened in Hubei, Shandong and Liaoning (see also Figure 1). Smoking has become significantly more prevalent in more rural areas. The higher rate of smoking in rural areas of China has been documented in other studies (He *et al*, 1995; Yang *et al*, 1999). The urbanicity index (thirds) itself has become more equally distributed, indicating that the most rural communities have been catching up in terms of infrastructure and services. Also the association between (log) income and urbanicity seems to have reduced by 2004, but one needs to be careful in comparing concentration indices for log incomes since the transformation will reduce the spread at the top of the distribution.

The associations between the covariates and obesity have remained quite stable over the time period. The effect of education is no longer significant in 2004. By 2004, Jiangsu and Hubei have reached higher obesity rates than Guizhou, whereas the difference between Guizhou and Hunan is no longer significant. Heavy physical activity now has a significant negative effect on obesity. The urbanicity index is still significantly positively associated with obesity.

The decomposition results of urbanicity-related inequalities in obesity in 2004 (see fourth and fifth column in Table 7) confirm the importance of the urbanicity index itself, which explains about 45% of the inequality. After this direct effect, obesity is lower in less urbanized areas because of the higher engagement in farming, which explains 44% of the inequality. Differences in physical activity account for a bit more than a quarter of urbanicity-related inequalities (27%). The next largest contribution comes from the higher incomes of more urbanized areas (11%). As in 1991, differences in education reduce inequality (-6%), but their contribution is no longer significant. Also the contribution of the demographics has disappeared, due to the lower age of the urban population as compared to 1991. Although jointly still insignificant, the province dummies now decrease urbanicity related inequalities in obesity (-4%). This is mainly due to the fact that the more southern provinces (Guangxi and Hunan) have become relatively more urbanized without experiencing a higher increase in obesity rates than the more northern and central regions (Liaoning, Shandong and Hubei) which have become relatively less urbanized, while maintaining their higher obesity rates.

The sixth column of Table 7 shows that quite a few of the associations between the covariates and hypertension have disappeared by 2004. Income, alcohol intake and the occupational variables are no longer significantly associated with hypertension. The urbanicity index itself, on the other hand, has become much more associated with hypertension, which suggests that community level factors have increased in importance relative to individual characteristics as determinants of hypertension.

The decomposition results are shown in the eighth and ninth column of Table 7. Due to the stronger association between hypertension and the urbanicity index, the direct contribution of the latter has increased to 62%. On the other hand, the contributions of demographics, income, province effects, and physical activity have all decreased. Presumably, the selective migration process discussed before is responsible for the decrease in the demographics contribution. The drop in the contribution of the province effects results from the same processes as discussed before.

Discussion

This paper offers a new methodology to quantify, track and explain the distribution of disease risk factors across areas at various stages of the urbanization process. This is done by constructing an urbanicity index using factor analysis on a broad set of community variables related to infrastructure, services and the local economy. The index outperforms the typical urban/rural dichotomy by detecting different degrees of urbanicity, measuring changes in urbanicity over time and being free from misclassification bias. By ranking communities on the basis of the index, we were able to measures of the degree of inequality in obesity and hypertension in relation to urbanicity in China. The results indicate that both in 1991 and in 2004, obesity and hypertension are more prevalent among the more urbanized areas. However, while the prevalence rates of these conditions have almost doubled over the period 1991-2004, inequalities across areas at different stages of urbanization have narrowed.

We used decomposition analysis to identify the factors that are driving these inequalities and their narrowing trend. More than half of the urbanicity related inequality in obesity is attributable to the direct effect of the community level characteristics that underlie the urbanicity index. This suggests that environmental factors are important determinants of obesity and that the higher prevalence in more urbanized areas is not simply an artifact of the characteristics of the population located there. For hypertension, the relative contribution of community level characteristics to the explanation of urbanicity related inequality more than doubles from 20% in 1991 to 62% in 2004. Environmental factors are becoming increasingly important determinants of the spread of this health condition. These could relate to changes in community diet and eating habits associated with the move away from agriculture and the rising number of restaurants. Improved transportation facilities, and increased use of motorized transport, contribute to more sedentary life styles and thereby to obesity and hypertension (Monda *et al*, 2007). Urbanization also brings about changes in the social context of communities, which have been shown to increase stress and herewith blood pressure (Niakara *et al*, 2007). Within the conceptual framework adopted these community factors are intermediate characteristics that impact on obesity/hypertension through proximate determinants. In the absence of data that perfectly capture these proximate determinants, the community variables will still explain some of the variation in obesity/hypertension rates. The finding that these community level characteristics still contribute substantially to inequalities in obesity and hypertension – even after controlling for the proximate determinants – does suggest that there is scope for community-level interventions to curb increasing obesity and hypertension prevalence rates.

Among the individual level factors, it is especially differences in engagement in physical activity and farming, both protective against obesity and hypertension and concentrated in more rural areas, that contribute to the lower prevalence of these health problems in less urbanized areas. This is in line with previous work which has argued that moving away from agriculture based employment to more sedentary work plays an important role in the rise of obesity in China (Monda *et al*, 2007). While both income and education are typically higher in urban areas, the former is associated with higher, and the latter with lower, obesity and hypertension. For this reason, the distribution of income increases and that of education decreases urbanicity-related inequalities in these conditions.

Three main factors appear to be associated with the declining urbanicity-related inequalities in both obesity and hypertension from 1991 to 2004. First, urban areas have obtained a higher concentration of young people, who are less prone to being obese or hypertensive. This is likely to result from the age-selective rural to urban migration and actually masks the urban concentration of these risk factors. Secondly, we find that those provinces that urbanized faster during the period 1991-

2004 (Hunan and Guangxi) did not experience higher increases in obesity/hypertension rates than provinces like Shandong, Liaoning and Hubei that were already relatively urbanized in 1991. This suggests that the more recent urbanization trends are less accompanied by rising obesity and hypertension. A third trend is related to the spread of urbanization itself. Our urbanicity index has become more equally distributed over time, which indicates that relatively more rural communities are catching up in terms of transport infrastructure, economic activity and community services. In other words, much of China is becoming more urbanized to some degree and with this the environmental conditions that encourage the spread of health problems such as obesity and hypertension are being established in *relatively* less urbanized communities.

An apparent fall in absolute differences in urbanicity seems to contrast with the huge media coverage given to the urban-rural divide in China (e.g. UN, 2008) and must be interpreted subject to two caveats. First, as discussed before, the CHNS data does not include the most western areas of China, where urban-rural inequalities have been shown to be the largest (Sicular *et al*, 2006), nor Shanghai and Beijing, two of the largest urban areas in the world. Therefore our estimates of urbanicity-related health inequalities do not necessarily reflect inequality trends for China as a whole, although they are based on the most consistent longitudinal data available with detailed measures of health and its determinants at both the individual and community level. Second, our composite index is estimated on a broad set of characteristics, and picks up well variations within urban and rural areas. However, it might have difficulty discriminating between the urbanized and the most urbanized, including megacities, as these are characterized not only by the community variables in our index, but also by other aspects such as modernization, technological progress, foreign investment etc.

In sum, this paper confirms that chronic health conditions associated with modernization and affluence, such as obesity and hypertension, are becoming a pressing problem in China, and, more originally, it reveals that the phenomenon is no longer an exclusively urban one. As development and urbanization are spreading within the Eastern and Central provinces of China, so are the diseases of affluence. Over the past 25 years, China has made extraordinary progress in reducing the number of people living in poverty, helping to combat its associated health problems. However, new chronic diseases are threatening some of the health gains from this progress. Given that universal health insurance coverage is still a long way off and consequently there is heavy reliance on direct payments for medical care (Liu *et al*; 1999), onset of a chronic illness represents a huge economic burden for millions of Chinese households. An important challenge lies in accompanying continued growth, development and urbanization with early preventive warnings that changing lifestyles will pose new health threats that ultimately carry their own economic costs.

References

Allen, W. (1976). *The Conceptualization and Measurement of Urbanization.* Dissertation, the University of Texas at Austin.

Champion, A.G., Hugo, G. (2004). *New forms of urbanization: Beyond the urban–rural dichotomy.* Aldershot, Hants, England, Burlington, VT: Ashgate.

Chou, S.Y., Grossman, M., Saffer, H. (2004). An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics* **23**: 565-587.

Dahly, L.D., Adair, L.S. (2007). Quantifying the urban environment: a scale measure of urbanicity outperforms the urban-rural dichotomy. *Social Science and Medicine* **64**: 1407-1419.

DeVellis, R.F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: SAGE Publications.

Gu, D.F., Reynolds, K., Wu, X., Chen, J., Duan, X., Reynolds, R.F., Whelton, P.K., He, J. and for the InterASIA Collaborative Group. (2005). Prevalence of the metabolic syndrome and overweight among adults in China. *Lancet* **365**: 1398-1405.

He, J., Klag, M.J., Wu, Z., Whelton, P.K. (1995). Stroke in the People's Republic of China. Geographic variations in incidence and risk factors. *Stroke* **26**: 2222-2227.

Heilig, G.K. (1999). *ChinaFood. Can China Feed Itself?* IIASA, Laxenburg (CD-ROM Vers. 1.1). Available at: http://www.iiasa.ac.at/Research/LUC/ChinaFood/data/urban/urban_8.htm

Hou, X. (2008). Urban-rural disparity of overweight, hypertension, undiagnosed hypertension, and untreated hypertension in China. *Asia-Pacific Journal of Public Health* **20**: 159-169.

International Diabetes Institute. (2002). *The Asia-Pacific perspective: Redefining obesity and its treatment.* Australia: Health Communications Australia Pty Limited.

Kim, S., Symons, M., Popkin, B. (2004). Contrasting socioeconomic profiles related to healthier lifestyles in China and the United States. *American Journal of Epidemiology* **159**: 184-191.

Kojima, R. (1995). Urbanization in China. *The Developing Economies* XXXIII-2.

Liu, L., Ikeda, K., Chen, M., Yin, W., Miushima, S., Miki, T., Nara, Y., Yamori, Y. (2004). Obesity, emerging risk in China: trend of increasing prevalence of obesity and its association with hypertension and hypercholesterolsemia among the Chinese. *Clinical and Experimental Pharmacology and Physiology* **31**: S8-S10.

Liu, Y., Hsiao, W.C., Eggleston, K. (1999). Equity in health and health care: the Chinese experience. *Social Science and Medicine* **49**: 1349-1356.

Liu, G.G., Dow, W.H., Fu, A.Z., Akin, J., Lance, P. (2008). Income productivity in China: on the role of health. *Journal of Health Economics* **27**: 27-44.

Liu, G.G., Wu, X., Chaoyang, P., Fu, A.Z. (2003). Urbanization and health care in rural China. Contemporary Economic Policy, **21**: 11-24.

Lukman, H., Dye, L., Blundell, J.E. (1993). Relationship between diet and obesity in Chinese groups. *British Food Journal* **100/1**: 3-9.

McDade, T.W., Adair, L.S. (2001). Defining the "urban" in urbanization and health: a factor analysis approach. *Social Science and Medicine* **53**: 55-70.

Monda, K.L., Adair, L.S., Zhai, F., Popkin , B.M. (2007). Longitudinal relationships between occupational and domestic physical activity patterns and body weight in China. *European Journal of Clinical Nutrition*: doi: 10.1038/sj.ejcn.1602849.

Monda, K.L., Gordon-Larsen, P., Stevens, J., Popkin, B.M. (2007). China's transition: The effects of rapid social change on adult activity patterns and overweight. *Social Science and Medicine* **64**: 858-870.

Monteiro, C.A., Moura, E.C., Conde, W.L., Popkin, B.M. (2004). Socioeconomic status and obesity in adult populations of developing countries: a review. *Bulletin of the World Health Organization* **82**: 940-946.

Morris, S., Sutton, M., Gravelle, H. (2003). Inequity and inequality in the use of health care in England: an empirical investigation. The University of York, Centre for Health Economics, Technical Paper Series 27.

National Bureau of Statistics China. (2000). Available at http://www.stats.gov.cn/english/

Niakara, A., Fournet, F., Gary, J., Harang, M., Nébié, L.V.A., Salem, G. (2007). Hypertension, urbanization, social and spatial disparities: a cross-sectional population based survey in a West African urban environment (Ouagadougou, Burkina Faso). *Transactions of the Royal Society of Tropical Medicine and Hygiene* **101**: 1136-1142.

Northridge, M.E., Sclar, E.D., Biswas, P. (2003). Sorting out the connections between the built environment and health: a conceptual framework for navigating pathways and planning healthy cities. *Journal of Urban Health* **80**: 556-568.

O'Donnell, O., van Doorslaer, E., Wagstaff, A. (2006). Decomposition of inequalities in health and health care, In: AM Jones (Ed.), *The Elgar Companion to Health Economics*, Edward Elgar, Cheltenham, chapter 17, 179-192.

Pan, Z. (2004). Socioeconomic predictors of smoking and smoking frequency in urban China: evidence of smoking as a social function. *Health Promotion International* **19**, 3, 309-315.

Popkin, B.M. (2003). The nutrition transition in the developing world. *Development Policy Review* **21**: 581-597.

Popkin, B.M., Du, S. (2003). Dynamics of the nutrition transition toward the animal foods sector in China and its implications: a worried perspective. *J. Nutr* **133**: 3898S-3906.

Reynolds, K., Gu, D., Muntner, P., Wu, X., Chen, J.; Huang, G., Duan, X., Whelton, P.K., He, J. for the InterASIA Collaborative Group. (2003). *Journal of Hypertension* **21**: 1273-1281.

Reynolds, K., Gu, D., Whelton, P., Wu, X., Duan, X., Mo, J., He, J. (2007). Prevalence and risk factors of overweight and obesity in China. *Obesity Research* **15**: 10-18.

Sahn, D.E., Stifel, D.C. (2000). Poverty comparisons over time and across countries in Africa. *World Development* **28**: 2123-2155.

Shen, J. (2006). Estimating urbanization levels in Chinese provinces in 1982-2000, *International Statistical Review* **74**: 89-107.

Sicular, T., Yue, X., Gustafsson, B., Li, S. (2006). The urban-rural income gap and inequality in China. United Nations University-World Institute for Development, Research paper 2006/135.

Sixth report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure. (1997). *Arch Intern Med* **157**: 2413–2446.

Ueshima, H., Zhang, X.H., Choudhury, S.R. (2000). Epidemiology of hypertension in China and Japan. *Journal of Human Hypertension* **14**: 765-769.

United Nations (UN). (2008). *Press conference by United Nations University to launch new study on Poverty, Growth, Inequality in China*. Available at: http://www.un.org/News/briefings/docs/2008/080605_China_Poverty.doc.htm

van Doorslaer, E., Gerdtham, U.G. (2003). Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data. *Social Science and Medicine* 57: 1621-1629.

van Doorslaer, E., Koolman, X., Jones, A. (2004). Explaining income-related inequalities in doctor utilization in Europe. *Health Econ* **13**: 629-647.

Vlahov, D., Galea, S. (2002). Urbanization, urbanicity, and health. *Journal of Urban Health* **79**: S1-S12.

Wagstaff, A. (2004). The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics* **14**: 429–432.

Wagstaff, A., Paci, P., van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social Science and Medicine* **33**: 545–7.

Wagstaff, A., van Doorslaer, E., Watanabe, N. (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics* **112**: 207-223.

Wang, Y. (2001). Cross-national comparison of childhood obesity: the epidemic and the relationship between obesity and socioeconomic status. *International Journal of Epidemiology* **30:** 1129-1136.

Wang, Y., Chen, J., Wang, K., Edwards, C.L. (2006). Education as an important risk factor for the prevalence of hypertension and elevated blood pressure in Chinese men and women. *Journal of Human Hypertension* **20**: 898-900.

Wang, L., Kong, L., Wu, F., Bai, Y., Burton, R. (2005). Preventing chronic diseases in China. *Lancet* **366**: 1821-1824.

Wang, Y., Mi, J., Shan, X.Y., Wang, Q.J, Ge, K.Y. (2007). Is China facing an obesity epidemic and the consequences? The trends in obesity and chronic disease in China. *International Journal of Obesity* **31**: 177-188.

Weng, X., Liu, Y., Ma, J., Wang, W., Yang, G., Caballero, B. (2007). An urban–rural comparison of the prevalence of the metabolic syndrome in Eastern China. *Public Health Nutrition* **10**:131-136.

Wooldridge, J.M. (2002). *Econometric analysis of Cross Section and Panel Data.* MIT Press: Cambridge.

World Bank. (2007). World Development Indicators 2007. The World Bank, Washington DC.

Wu, H.X. (1994). Rural to urban migration in the People's Republic of China. *The China Quarterly* **139**: 669-698.

Wu, X. Duan, X., Gu, D., Hao, J., Tao, S., Fan, D. (1995). Prevalence of hypertension and its trends in Chinese populations. *International Journal of Cardiology* **52**: 39-44.

Yan, X. (1990). The spatial dimension of Chinese Urbanization. *GeoJournal* **21.1/2**: 13-23.

Yang, G., Fan, L., Tan, J., Qi, G., Zhang, Y., Samet, J., Taylor, C.E., Becker, K., Xu, J. (1999). Smoking in China. Findings of the 1996 National Prevalence Survey. *JAMA* **282**: 1247-1253.

Yusuf, F., Saich, T. (2008). *China urbanizes: consequences, strategies and policies*. The World Bank, Washington DC.

Zhao, L., Stamler, J., Yan, L.L., Zhou, B., Wu, Y., Liu, K., Daviglus, M.L., Dennis, B.H., Elliott, P., Ueshima, H., Yang, J., Zhu, L., Guo, D. for the INTERMAP Research Group. (2004). Blood pressure differences between northern and southern Chinese: role of dietary factors: the International Study on Macronutrients and Blood Pressure. *Hypertension* **46**: 1332-1337.

Tables and Figures

Table 1: Description of community characteristics used in the estimation of the urbanicity index, classified according to the framework proposed by Northridge *et al* (2003).

built environment	land use	farmland	whether there is any farmland in the comm (1-0)
		workagri	% of workforce in comm that is working in agriculture (%)
	transportation	bus	whether comm is near a bus station (1-0)
		train	whether comm is near a train station (1-0)
		road1	whether dirt is main characteristic of roads in comm (1-0)
		road2	whether stone/gravel is main characteristic of roads in comm (1-0)
		road3	whether paved is main characteristic of roads in comm (1-0)
		paved	whether there is any paved road in comm (1-0)
		pavedkm	distance to nearest paved road from comm (km)
	services	allmarketkm	average distance to markets (average over different goods) (km)
		telephone	whether comm has convenient telephone service (1-0)
		post	whether there is a post office in comm (1-0)
		newspaper	whether comm can receive daily newspaper on the day it is published (1-0)
		prim	whether there is a primary school in the comm (1-0)
		mid	whether there is a secondary school in de comm (1-0)
		voc	whether there is a vocational school in the comm (1-0)
		meankmfac	average distance across all health facilities people in comm can go to (km)
		cutoff	average number of days (a week) that electricity is cut off in comm (days)
		childcare3	wether there is a child care center (children<3) in comm (1-0)
		childcare6	whether there is a child care center (children<6) in comm (1-0)
		resto	number of restaurants in comm
social-economic context		enterprises	number of enterprises in comm
		bigent	% of workforce that is working in enterprises with >20 people (%)
		smallent	% of workforce that is working in enterprises with <20 people (%)
		ota	whether there is an open trade area in comm (1-0)
		popul	population of comm

Table 2: Description of proximate variables.

demographics	age-gender dummies	M1629 (males aged between 16 and 29), M3044, M4564, M65plus F1629 (females aged between 16 and 29), F3044, F4564, F65plus
	married	whether resondent is married (1-0)
lifestyle	activity	whether respondent performs (very) heavy physical activity in daily activities (1-0)
	logfat	logaritm of average daily gramms of fat intake
	smoke	whether respondent is currently smoking (1-0)
	alcodaily	whether respondent consumes alcohol on a daily basis (1-0)
socioeconomic status	noeduc	whether respondent has had no education (1-0)
	edprim	whether respondent's highest education is primary education (1-0)
	edmid	whether respondent's highest education is secondary education (1-0)
	edhigh	whether respondent's highest education is higher education (1-0)
	logincome	logarithm of household income (in Chinese Yuan)
	notwork	whether respondent is not working (1-0)
	professional	whether respondent is in a professional occupation (1-0)
	farmer	whether respondent is involved infarming (1-0)
	skilwork	whether resondent is a skilled worker (1-0)
	nonskilwork	whether respondent is a non-skilled worker (1-0)
	otherwork	whether respondent is invloved in any other occupation (1-0)

Table 3: Summary statistics of intermediate (community) variables across waves.

Variable	1991	1993	1997	2000	2004
farmland	0.61	0.60	0.59	0.52	0.55
workagri	47.27	42.70	41.30	39.24	32.95
bus	0.55	0.54	0.62	0.69	0.63
train	0.19	0.19	0.17	0.19	0.23
road1	0.27	0.23	0.21	0.15	0.07
road2	0.23	0.23	0.21	0.19	0.23
road3	0.50	0.55	0.58	0.66	0.70
paved	0.86	0.88	0.89	0.93	0.98
pavedkm	0.61	0.27	0.32	0.08	0.03
allmarketkm	1.03	0.86	0.80	0.23	0.22
telephone	0.56	0.66	0.81	0.93	0.89
post	0.84	0.88	0.88	0.88	0.82
newspaper	0.31	0.39	0.49	0.46	0.59
prim	0.66	0.73	0.72	0.77	0.72
mid	0.29	0.32	0.33	0.32	0.35
voc	0.07	0.09	0.09	0.09	0.09
meankmfac	5.35	3.22	4.62	4.10	2.90
cutoff	1.20	0.87	0.60	0.47	0.31
childcare3	0.20	0.21	0.28	0.24	0.32
childcare6	0.45	0.48	0.54	0.43	0.54
resto	5.04	6.23	7.10	8.98	10.95
enterprises	36.73	32.38	40.23	97.95	174.81
bigent	32.08	29.80	27.61	30.40	31.33
smallent	8.41	10.38	15.99	18.35	17.09
ota	0.22	0.41	0.38	0.43	0.52
popul	2247.98	2746.67	2256.70	3524.06	4932.85
Observations	189	181	167	191	192

Table 4: Factor loading for intermediate (community) variables

variable	factor loading
farmland	-0.65
workagri	-0.73
bus	0.30
train	0.23
road1	-0.62
road2	-0.27
road3	0.72
paved	0.45
pavedkm	-0.21
allmarketkm	-0.33
telephone	0.45
post	0.20
newspaper	0.46
prim	-0.03
mid	0.21
voc	0.25
meankmfac	-0.24
cutoff	-0.27
childcare3	0.43
childcare6	0.33
resto	0.47
enterprises	0.37
bigent	0.46
smallent	0.28
ota	0.27
popul	0.29

NOTE: Factor analysis is used upon data pooled across all waves (1991, 1993, 1997, 2000 and 2004).

Table 5: Averages of obesity, hypertension, fat intake and income across quintiles of the urbanicity index

	Q1	Q2	Q3	Q4	Q5
obesity	0.10	0.12	0.17	0.25	0.26
hypertension	0.12	0.14	0.16	0.21	0.25
fat intake (grams)	58.18	64.95	67.10	75.90	81.84
income (Yuan)	1526	2293	2842	3431	4587

NOTE: Table based upon data pooled across all waves.

Table 6: Decomposition of urbanicity related inequalities in obesity and hypertension in 1991.

		Obesity				Hypertension			
	concentration index	coefficient	elasticity	% contribution	% contribution	coefficient	elasticity	% contribution	% contribution
M3044	-0.032*	0.057***	0.066	-0.009*	0.050***	0.051***	0.056	-0.010	0.280***
M4564	0.027*	0.079***	0.074	0.008		0.186***	0.161	0.023	
M65plus	0.248***	0.056**	0.015	0.015**		0.369***	0.091	0.121***	
F1629	-0.042***	-0.014	-0.020	0.003		-0.059***	-0.077	0.017***	
F3044	-0.046***	0.082***	0.113	-0.021***		-0.007	-0.010	0.002	
F4564	0.059***	0.134***	0.136	0.033***		0.121***	0.114	0.036***	
F65plus	0.188***	0.088***	0.027	0.021***		0.316***	0.089	0.089***	
logincome	0.035***	0.014***	0.748	0.110***	0.110***	0.013***	0.656	0.125***	0.125***
married	0.002	0.035***	0.199	0.002	0.002	0.001	0.008	0.000	0.000
edprim	-0.135***	0.010	0.016	-0.009	-0.101***	-0.022*	-0.033	0.024*	-0.110***
edmid	0.046***	-0.018	-0.040	-0.007		-0.045***	-0.092	-0.022***	
edhigh	0.317***	-0.053***	-0.065	-0.085***		-0.057***	-0.066	-0.111	
Jiangsu	-0.013	-0.008	-0.008	0.000	0.038	-0.014	-0.013	0.001	0.131***
Shandong	0.216***	0.100***	0.100	0.090***		0.082***	0.077	0.089***	
Henan	-0.253***	0.098***	0.077	-0.081***		0.057***	0.042	-0.057***	
Hubei	0.064**	0.012	0.012	0.003		0.069***	0.065	0.022*	
Hunan	-0.020	-0.041***	-0.037	0.003		0.031**	0.026	-0.003	
Guangxi	-0.007	-0.067***	-0.052	0.001		-0.016	-0.012	0.000	
Liaoning	0.117***	0.047***	0.043	0.021**		0.147***	0.125	0.078***	
activity	-0.352***	-0.012	-0.046	0.067	0.067	-0.044***	-0.153	0.289***	0.289***
alcodaily	0.038**	0.015	0.011	0.002	0.002	0.078***	0.052	0.010*	0.010*
logfat	0.015***	0.023***	0.695	0.043***	0.043***	0.003	0.086	0.007	0.007
smoke	-0.0120	-0.050***	-0.121	0.006	0.006	-0.036***	-0.081	0.005	0.005
farmer	-0.398***	-0.071***	-0.244	0.403***	0.288***	0.014	0.047	-0.100	0.041
skilwork	0.371***	-0.024	-0.013	-0.019		-0.012	-0.006	-0.012	
nonskilwork	0.318***	-0.064***	-0.041	-0.054***		-0.004	-0.003	-0.004	
otherwork	0.356***	-0.052***	-0.030	-0.045**		-0.011	-0.006	-0.012	
notwork	0.291***	0.002	0.002	0.003		0.076***	0.108	0.169***	
urbmid	0.167***	0.009	0.020	0.014	0.503***	-0.003	-0.005	-0.005	0.194*
urbhigh	0.726***	0.079***	0.163	0.489***		0.026*	0.051	0.198**	
error	-0.002			-0.009	-0.009			0.028	0.028
total				1.000	1.000			1.000	1.000

NOTES: Coefficients from linear probability model. * indicates significance at 10%, ** at 5%, *** at 1% (based upon bootstrapped standard errors).

Table 7: Decomposition of urbanicity related inequalities in obesity and hypertension in 2004.

		Obesity				Hypertension			
	concentration index	coefficient	elasticity	% contribution	% contribution	coefficient	elasticity	% contribution	% contribution
M3044	-0.010	0.092***	0.045	-0.005	-0.032	0.060***	0.035	-0.003	0.168***
M4564	-0.034***	0.109***	0.082	-0.030*		0.229***	0.200	-0.067*	
M65plus	0.101***	0.057**	0.014	0.016		0.355***	0.106	0.105***	
F1629	0.049*	-0.074***	-0.019	-0.010		-0.084***	-0.026	-0.012	
F3044	-0.051***	0.042*	0.025	-0.014		-0.034*	-0.024	0.012	
F4564	-0.018	0.137***	0.114	-0.022		0.114***	0.111	-0.019	
F65plus	0.133***	0.080**	0.023	0.033**		0.351***	0.116	0.153***	
logincome	0.024***	0.014**	0.438	0.114**	0.114**	0.003	0.103	0.025	0.025
married	-0.011***	0.070***	0.213	-0.027**	-0.027**	0.011	0.038	-0.004	-0.004
edprim	-0.145***	0.001	0.000	-0.001	-0.056	-0.0120	-0.012	0.017	-0.113**
edmid	-0.022*	-0.012	-0.014	0.003		-0.037**	-0.052	0.011	
edhigh	0.285***	-0.021	-0.019	-0.059		-0.047**	-0.050	-0.142**	
Jiangsu	0.006	0.092***	0.056	0.003	-0.035	0.078***	0.055	0.003	-0.073*
Shandong	0.075***	0.209***	0.102	0.082***		0.089***	0.051	0.038**	
Henan	-0.080***	0.173***	0.075	-0.065**		0.094***	0.048	-0.038*	
Hubei	-0.194***	0.049**	0.020	-0.042**		0.098***	0.047	-0.091***	
Hunan	0.157***	-0.004	-0.001	-0.002		0.010	0.004	0.006	
Guangxi	0.062**	-0.064***	-0.030	-0.020		-0.001	0.000	-0.000	
Liaoning	0.012	0.171***	0.078	0.010		0.146***	0.078	0.009	
activity	-0.408***	-0.045***	-0.061	0.267**	0.267**	-0.026*	-0.041	0.166*	0.166*
alcodaily	-0.019	-0.021	-0.009	0.002	0.002	0.023	0.011	-0.002	-0.002
logfat	0.015***	0.027**	0.426	0.067**	0.067**	-0.003	-0.065	-0.009	-0.009
smoke	-0.034***	-0.062***	-0.065	0.024***	0.024***	-0.040**	-0.049	0.016*	0.016*
farmer	-0.441***	-0.080***	-0.092	0.439***	0.254**	-0.019	-0.026	0.111	0.153*
skilwork	0.243***	-0.060	-0.009	-0.023		0.018	0.003	0.007	
nonskilwork	-0.001	-0.067**	-0.014	0.000		-0.007	-0.002	0.000	
otherwork	0.256***	-0.046*	-0.018	-0.049*		-0.005	-0.002	-0.006	
notwork	0.196***	-0.037	-0.054	-0.113		0.012	0.021	0.041	
urbmid	-0.276***	0.012	0.017	-0.050	0.444***	0.021	0.035	-0.095	0.619***
urbhigh	0.551***	0.049***	0.083	0.495***		0.066***	0.131	0.714***	
error	0.005			-0.022	-0.022			0.053	0.053
total				1.000	1.000			1.000	1.000

NOTES: Coefficients from linear probability model.* indicates significance at 10%, ** at 5%, *** at 1% (based upon bootstrapped standard errors).

Figure 1: Map of CHNS provinces, colored according to their average urbanicity index in 1991 and 2004.

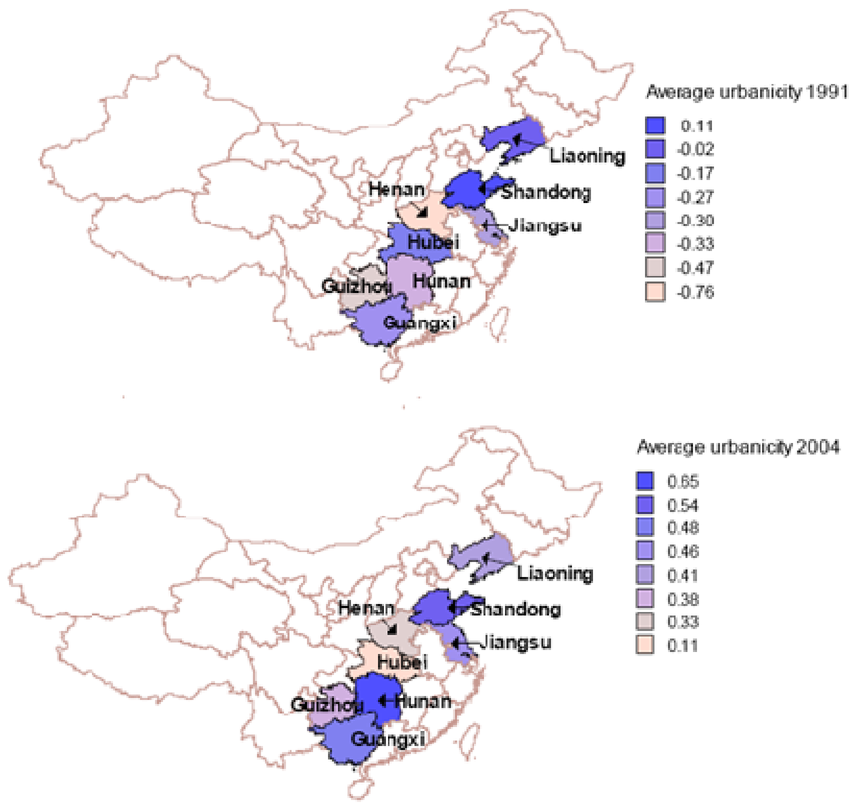
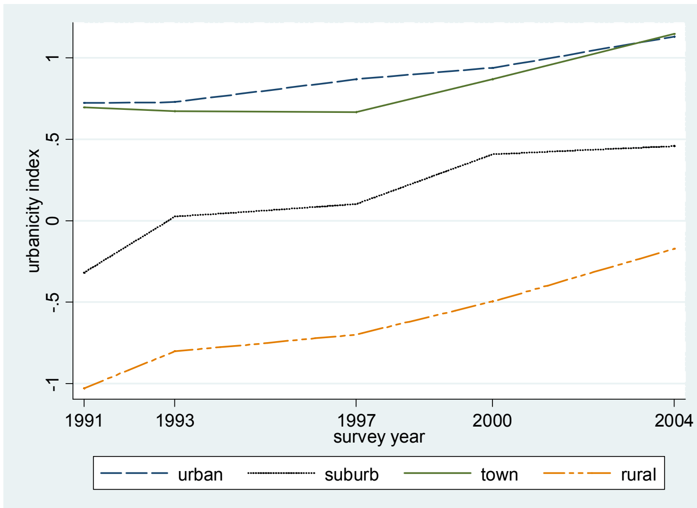
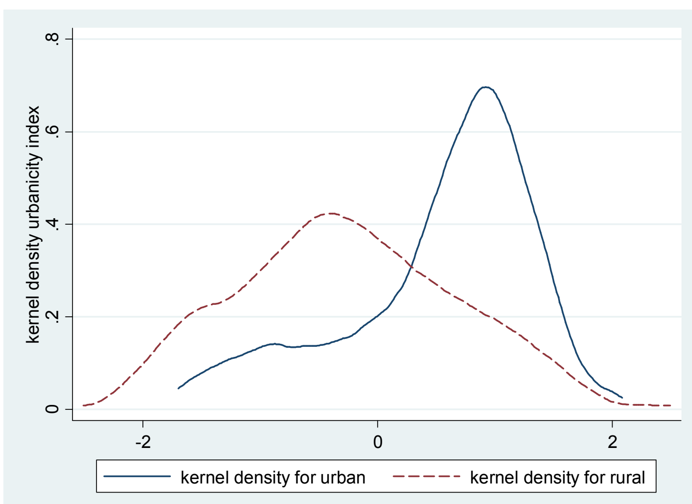


Figure 2: Time trends in urbanicity index across the defined set of settlement classifications as available in the CHNS.



NOTE: Based upon data pooled across all waves (1991, 1993, 1997, 2000 and 2004).

Figure 3: Kernel density estimates of the urbanicity index for urban/rural communities as defined in the CHNS.



NOTE: Based upon data pooled across all waves (1991, 1993, 1997, 2000 and 2004).

Notes

[1] Following McDade and Adair (2001) and Vlahov and Galea (2002), we use the term ''urbanization'' to describe the process by which communities become increasingly urban and the term ''urbanicity'' to describe the degree to which a community has the characteristics of an urban environment. Urbanization is a process, whereas urbanicity is a state at any point in time in that process.
[2] In the 1989 survey, health and nutritional data were only collected from preschoolers and adults aged 20-45.
[3] We have not included Heilongjiang in our analysis as this province only joined the CHNS in 1997.
[4] Note that we do not exploit the panel nature of the data as a balanced panel would not be representative and in turn, the urbanicity related inequalities would only be meaningful for individuals in the panel.
[5] Results were robust to not taking into account whether individuals were taking blood pressure lowering medication.
[6] This lower cut-off has been recommended by the International Obesity Task Force for defining obesity in Asian populations (International Diabetes Institute, 2000).
[7] To include as much information as possible in the estimation of the urbanicity index, we use data of all waves (1991, 1993, 1997, 2000 and 2004). Information on the size of the community (area) is not collected in 1997 and 2000. In 2004, information on the size of the community is gathered retrospectively for all previous waves (so only for the balanced panel). In the pooled sample, for 15% of the communities we could not calculate population density. As individuals are nested within communities, there is a much higher percentage of missing information on density in the individual level data (22%).
[8] Nutrient intake variables are calculated by the CHNS team and based upon detailed food intake information regarding three days before the survey. We chose to use fat intake as this is more reflecting dietary preferences then calorie intake (Lukman, Dye and Blundell, 1998).
[9] For more information see http://www.cpc.unc.edu/projects/china/data/datasets/convar.html. This is the income measure (and deflator) that is used by Liu *et al* (2008). We also checked robustness of the results when using GDP deflators for all China from the World Development Indicators (World Bank, 2007).
[10] Another option would be to use principal component analysis, which assumes that *all* variability in an item should be used in the analysis, while in principal factors analysis we only use the variability in an item that it has in common with the other items. Principal components analysis is often preferred as a method for data reduction, while principal factors analysis is often preferred when the goal of the analysis is to detect structure. Sahn and Stifel (2000) argue that factor analysis is more suited if one wishes to extract only 1 factor. In most cases, these two methods usually yield very similar results, and this was also confirmed in our study
[11] Note that the urban-rural categorization in the CHNS does not vary across waves.
[12] Note that the normalization consists of dividing the standard concentration index by (1-H).
[13] It is important to note that, as opposed to income related inequalities, urbanicity related inequalities are not necessarily undesirable if these would result from purely geographical reasons
[14] In the decomposition of the standard concentration index, a variable's contribution to inequality comes from the product of its elasticity and standard concentration index, while in the decomposition of the normalized concentration index these contributions are scaled by the factor (1-H).
[15] As our data consists of individuals nested within communities, we could use community random effects models to explain obesity/hypertension. This would then allow estimating the community random components and their contribution to inequality. However, as these random components would be assumed independent of all the covariates (including the urbanicity index), their concentration index should be close to zero. We have confirmed this empirically.
[16] The proportion of farmers decreases steadily across quintiles of the urbanicity index, but even in the upper quintile 2% of individuals are involved in farming.