



Published in final edited form as:

*Soc Sci Med.* 2010 October ; 71(8): 1436–1446. doi:10.1016/j.socscimed.2010.07.027.

## Understanding Community Context and Adult Health Changes in China: Development of an Urbanicity Scale

Jessica Jones-Smith and Barry M. Popkin

### Abstract

The classification of places as either urban or rural is typically based on an absolute threshold of population and/or population density. However, conceptual definitions of urbanization and urbanicity encompass dimensions beyond solely population size and population density. Multiple important distinguishing urban characteristics beyond population size have been described. The crude classification of places as urban or rural coupled with infrequent updates to this information create a measure that is prone to misclassification error. An improved measure of urbanicity would draw information from the domains that characterize urban and rural places, would be sensitive to changes over time, and would represent gradations on the continuum from rural to urban environments. The goal of the current study was to develop such a scale from existing data, test whether the scale was reliable and valid, and assess whether it provided information beyond what could be determined from the traditional urban/rural dichotomous variable. We utilized established scaling procedures from the psychometric literature to construct and evaluate a multicomponent scale to measure urban features on a continuum in China. We also provided an example of its potential contribution to health research by examining its relationship with the adult body mass index (BMI). Because the scale was constructed and tested using established scaling procedures and using a wide array of variables, it represents an improvement over previous attempts at such a scale and will provide a reliable and valid measurement tool for researchers in this arena. The scale was developed to predict the incidence of overweight/obesity populations in China, but it promises to be most useful for other economic, demographic, social welfare, and health outcomes.

### Keywords

China; Urban/Rural; China Health and Nutrition Survey; Urbanicity Scale; Obesity; Overweight; BMI; methodology

## INTRODUCTION

The classification of places as either urban or rural is typically based on an absolute threshold of population and/or population density (Champion & Hugo, 2004). However, conceptual definitions of urbanization and urbanicity encompass dimensions beyond solely population size and population density (Champion & Hugo, 2004). Urbanization has been defined as the change in size, density, and heterogeneity in places and the migration into cities, while urbanicity has been described as the presence of urban features or the degree to which a place

---

Corresponding Author: Barry M. Popkin, University of North Carolina at Chapel Hill, Chapel Hill, NC UNITED STATES  
popkin@unc.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

exhibits urban characteristics (Vlahov & Galea, 2002). Multiple important distinguishing urban characteristics have been described and include: population and proximity, infrastructure, diversity, social environment, culture, and economic activity (Champion & Hugo, 2004; National Research Council, 2003; Vlahov & Galea, 2002). The standard operationalization of places as either urban or rural does not account for many of these key urban characteristics and cannot distinguish cities on a continuum of urbanization or urbanicity (Brockhoff & Brennan, 1998; Champion & Hugo, 2004; McDade & Adair, 2001; Mendez & Popkin, 2005).

Additionally, trajectories of urban development likely differ spatially, temporally, and by degree of baseline urbanicity. Many villages throughout the developing world have gained only some dimensions constituting an urban environment (Zhu, 2004), while other demographic entities are losing factories, but gaining universities, markets, and transportation infrastructure (Rondinelli, 1983, 1987). The crude classification of places based on arbitrary demographic thresholds coupled with infrequent updates to this information creates a measure that is quite prone to misclassification error and that becomes less accurate as the time since the last census increases (Champion & Hugo, 2004).

Despite the limitations of the urban-rural dichotomous classification, differences in human health according to the urbanicity of places have been widely documented (Brockhoff & Brennan, 1998; Fotso, 2006; Johansson & Mosk, 1987; Monda, Gordon-Larsen, Stevens, & Popkin, 2007; B. M. Popkin, Bilsborrow, & Akin, 1982; Preston & Walle, 1978; Van de Poel, O'Donnell, & Van Doorslaer, 2007; Vlahov & Galea, 2002). Recent studies on infant mortality, obesity, breastfeeding have all found that urban residence represents a critical dimension (Batal, Boulghourjian, Abdallah, & Afifi, 2006; Brockhoff & Brennan, 1998; Hinde & Mturi, 1996; Mendez, Monteiro, & Popkin, 2005; B. M. Popkin et al., 1982). However, variations in health by crude classification of urban/rural status likely do not capture the enormous heterogeneity that is emerging in rural and urban areas (Allender, Foster, Hutchinson, & Arambepola, 2008; McDade & Adair, 2001; Montgomery & Ezeh, 2005).

As theories continue to develop to explain the effects of rural and urban environments on health, there is a need for methods to provide a more refined measure of urbanicity based on its various dimensions (Champion & Hugo, 2004). This strategy would enable both an inter- and intra-urban and rural comparison (Champion & Hugo, 2004; Montgomery & Ezeh, 2005; National Research Council, 2003). An improved measure of urbanicity would draw information from the domains that characterize urban and rural places, be sensitive to changes over time, and represent gradations on the continuum from rural to urban environments.

The goal of the current study was to determine whether such a scale could be developed from existing data, whether the scale was reliable and valid, and whether it provided information beyond what could be determined from the traditional urban/rural dichotomous variable. We utilized scaling procedures from the psychometric literature to construct and evaluate a multicomponent scale to measure urban features on a continuum in China. We then briefly explored the relation between urbanicity and incident overweight/obesity status among adult women in China to demonstrate the utility of this new scale for investigating the relations between urbanicity and health outcomes.

First, we provide a short overview of some of the important features of the China Health and Nutrition Survey and of China's recent history in relation to urbanization processes. In addition, we compare our scale-building approach with other recently published urbanicity scales. Subsequently, we describe our methods and results and provide comment.

China provided a unique opportunity to develop and test our scale. Extensive community-level data have been collected from 1989 to 2006 in the China Health and Nutrition Survey (CHNS). The CHNS captured the impact of China's reforms, including: eliminating most food subsidies

and reducing the remaining food and fuel subsidies, allowing freely fluctuating market prices, permitting and creating conditions to establish a private sector in most areas of economic activity, and initiating processes that will ultimately transfer much property and land to private ownership. China began reforms in 1978, at a time of relative macroeconomic stability, with the majority of its labor force tied to agricultural activity. The reforms of the 1978–1980 period have been followed by many shifts in economic and social policies at the national and provincial levels and in many cases also at the city/county level. Tens of millions have migrated initially toward eastern coast provinces and more recently to cities and towns across the country (Friedman, 2005; Hu, Cook, & Salazar, 2008; Ma, 2002; Poston & Mao, 1998; Yang & Guo, 1999). Whereas urban development used to be highly controlled through funneling funds and resources based on the administrative status assigned to a city, town, or village, with increasing privatization and capitalistic practices, the urbanicity of a place in China is now influenced by economic and social contextual forces rather than previous administrative dictates.

The communities in the CHNS sample are drawn from cities, suburbs, towns, or villages of China, all entities that are legally identified by the National Bureau of Statistics of China. China classifies not only places, but also people, as either rural or urban (Shen, 1995). This classification has traditionally been based on a family's registration status (*Hukou*) as either part of the agricultural population or part of the nonagricultural population (Shen, 1995). The focus of this paper is on the measurement of the urbanicity of places and communities, rather than on the official urban/rural registration status of Chinese people.

To our knowledge, there are only a few published reports of scales that attempt to measure urbanicity on a continuum. McDade and Adair (2001) used a factor analysis approach to evaluate level of urbanicity for the Cebu Longitudinal Health and Nutrition Survey. Subsequently, Dahly and Adair (2007) developed and tested the reliability and validity of a new scale to measure urbanicity in Cebu, Philippines. The Cebu urbanicity scale was influential in our motivation to develop the scale to measure urbanicity in China. Similar to the methods used by Dahly and Adair (2007), we used published literature to guide our definition of urbanicity, which in turn guided our scale development. This sequence contrasts with, and has advantages over a data-derived factor analytic approach. Establishing a strong working definition of the underlying construct and its content domain is an integral step in scale development (DeVellis, 2003) that can be overlooked or underemphasized when relying on data-driven methods. We also conducted a content validity survey among CHNS senior investigators and staff to aid in the establishment of our context-specific content domain (DeVellis, 2003). Contrary to this approach, Van de Poel, O'Donnell & Van Doorslaer (2009) have developed a factor-analysis-derived urbanicity index using the CHNS data. The scale that we present herein uses more variables and encompasses more domains compared to the factor analysis-derived model in Van de Poel et al. (2009). Additionally, we developed our measure as a scale and used scaling procedures to test and validate our measure, whereas Van de Poel et al. (2009) performed limited tests of their tool's performance as a scale (DeVellis, 2003). Compared to the Cebu urbanicity scale, our scale for China includes more components of urbanicity, and its reach spans a much larger geographic area and population. Additionally, our scale revised the scoring algorithms of Dahly and Adair (2007) to provide finer gradations in component scores and to allow room for growth in future years of the survey.

## METHODS

The CHNS data collection began in 1989 and has been implemented every 2 to 4 years since. The CHNS uses a multistage cluster sample design to survey individuals and households within 218 neighborhoods within nine provinces in China. These nine provinces contain approximately 56% of the population of China. To obtain the sample from these nine provinces, the counties inside the provinces were stratified by income then a weighted sample of four

counties was selected, as was the provincial capital city and a lower-income city in each province. Within these cities or counties, neighborhoods were randomly selected, resulting, originally, in 190 communities and, currently, in 218 communities. The mean (standard deviation) number of households per community was 21 (2.1). Households were selected randomly from a community household roster and all members in each household were interviewed. The household roster was used to follow up each of the originally sampled households as well as new households formed from previous households for subsequent survey panels. The baseline sample was representative of each province but over time, loss-to-follow up has occurred. Overall, 85% of households have been surveyed in at least five of seven rounds (Barry M. Popkin, Du, Zhai, & Zhang, 2009). The CHNS includes individual, household, and community level surveys conducted by trained field worker; the current analysis utilizes information from each of these surveys. The household and individual surveys were conducted inside the participant's home. The community survey obtained detailed information on the community infrastructure, services, and demographic/economic environment from a combination of neighborhood officials, informants, visits to markets, and official records (Monda et al., 2007). Ethical approval was provided by the Internal Review Board of University of North Carolina at Chapel Hill and Chinese Center for Disease Control

### Scale Construction

**Overview**—Motivated by established procedures for building scales (DeVellis, 2003; Netemeyer, Bearden, & Sharma, 2003), we began by establishing a strong working definition of the construct that we intended our scale to measure. To do so, we consulted authoritative sources, previous works, and content experts. Next, we determined which variables were available to represent the concepts identified as important defining and distinguishing features and how each variable should be scored. Finally, we tested the scale's performance as a measurement tool, including its dimensionality, reliability, content, criterion and construct validity. Each of these procedures is described below.

**Variable Selection**—To determine which variables should be used to construct the urbanicity scale, we consulted references to find common defining and distinguishing features of urban places and from these we synthesized a list of features to focus our definition of urbanicity (Champion & Hugo, 2004; National Research Council, 2003; Vlahov & Galea, 2002). This included: population size, proximity (population density), type of occupations and percent employed in agriculture, number of markets, reliance on cash systems (as opposed to barter) in markets, diversity of markets, infrastructure (such as piped water, waste disposal, paved roads, communication systems, transportation, electricity), different social networks and culture, higher average education and income, and greater diversity in education and income. We used this list to identify existing CHNS survey items that could be used to operationalize these characteristics. Unfortunately, we could not find data for a few characteristics such as social networks and culture and so these are not included in the scale. We identified 12 broad areas, which we call components, to distinguish urbanicity. Multiple variables were used to construct some of these components, whereas other components are represented by a single variable.

After identifying potential components and variables to be used in our scale of urbanicity, we conducted a small (n=5) content validity survey to assess how well the collection of variables we assembled and the proposed scoring methods for each seemed to capture the content domain of urbanicity (DeVellis, 2003). These evaluators included senior members of the CHNS research team as well as former CHNS staff members with extensive experience in both rural and urban China. We collected comments and suggestions from these reviewers and incorporated them into the final version of the variables to be used and the scoring assignment for each variable.

**Variables Included**—We identified 12 components thought to define and distinguish urbanicity that could be operationalized in the CHNS data:

- Population density: total population of the community divided by community area, from official records.
- Economic activity: typical daily wage for ordinary male worker (reported by community official) and percent of the population engaged in nonagricultural work.
- Traditional markets: distance to the market (three categories), (1) within the boundaries of the community, (2) within the city but not in this community, or (3) not within the city/village/town); number of days of operation for eight different types of market (including food and fuel markets).
- Modern markets: number of supermarkets, cafes, internet cafes, indoor restaurants, outdoor fixed and mobile eateries, bakeries, ice cream parlors, fast food restaurants, fruit and vegetable stands, bars within the community boundaries.
- Transportation infrastructure: most common type of road, distance to bus stop, and distance to train stop. (Distance is categorized as (1) within community, (2)  $\leq 1$  km from community, and (3)  $> 1$  km from community).
- Sanitation: proportion of households with treated water and prevalence of households without excreta present outside the home.
- Communications: availability (within community boundaries) of a cinema, newspaper, postal service, telephone service; and percent of households with a computer, percent of households with a television, and percent of households with a cell phone.
- Housing: average number of days a week that electricity is available to the community, percent of community with indoor tap water, percent of community with flush toilets, and percent of community that cooks with gas.
- Education: average education level among adults  $> 21$  years old.
- Diversity: variation in community education level and variation in community income level.
- Health infrastructure: number and type of health facilities in or nearby (  $12$  km) the community and number of pharmacies in community.
- Social services: provision of preschool for children under 3 years old, availability of (offered in community) commercial medical insurance, free medical insurance, and/or insurance for women and children.

Variables described as percentages from households were derived from the household surveys in the communities. All other variables are collected as part of the community survey. We originally included the community's total population, but since our communities can consist of either an entire village or only a few blocks in a major city, the total population of the community is not sensitive or meaningful in this case as an urban indicator, so we excluded total population from the final version.

With a few exceptions, all variables were surveyed each year from 1991 to 2006. The first survey year, 1989, did not include the assessment of ordinary male wage. The variables in the social service component were added to the survey in 2000, and the survey of supermarket and other modern food vendors was added in 2004. Cell phone ownership was added in 2004, and computer ownership began in 1997. Additionally, the first survey year, 1989, had higher frequencies of missing observations than following years.

Observations were missing for 22% of the total data points in 1989, 11 to 13% for 1991–2000, and 4 to 5% for 2004 and 2006. For survey year 1989, missing observations were imputed based on the community's value in the 1991 survey. In subsequent survey years, imputations were based on the community's value in the preceding year.

**Additional Variables**—In the brief applications of the new scale, we utilized individual measured height and weight measurements. Body mass index (BMI) was calculated from weight in kilograms divided by height in meters squared and used to categorize individuals as overweight or obese ( $BMI \geq 25$ ) (WHO/FAO, 2003). We also briefly explored how urbanicity was associated with dietary changes by using the percentage of calories from fat as a dependent variable. Percentage of calories from fat was obtained from each individual with a 24-hour dietary recall conducted for 3 consecutive days. In addition to the 24-hour recall, a daily household inventory of use of all major food staples, including oils, was collected at the beginning and end of each day. All data sources were combined to measure each individual's total intake. We calculated the 3-day average percentage of calories from fat.

We also utilized inflation-adjusted per capita family income, which was derived from reports of income from multiple sources for all families. Additionally, we used the number of surviving children born to each married woman obtained from the individual survey.

### Scale Scoring

We allotted a maximum total of 10 points each to each of the 12 components. We had no *a priori* reason to justify weighting certain components more heavily than others, so we followed the scheme of Dahly and Adair (2007). Scoring algorithms within these components were developed based on distributions in the data, with the goal of having the median score in a middle year be close to half of the total possible points and with sufficient spread in the scores between the minimum and maximum points (DeVellis, 2003). The supplementary data section provided details on the scoring algorithm used for each component [LINK TO SUPPLEMENTARY DATA FILE]

### Scale Properties, Reliability and Validity

To evaluate whether the components operated at a unidimensional or a multidimensional scale, we first performed an exploratory factor analysis on the 12 components, without restricting the number of factors estimated. We assessed dimensionality by the number of factors with eigenvalues  $>1$  and with a scree plot (Netemeyer et al., 2003). After unidimensionality for one underlying construct was demonstrated, we assessed the internal consistency of the scale using Cronbach's alpha, which is an estimate of the shared covariance among the scale components (DeVellis, 2003). Additionally, corrected item-scale correlations were reported (DeVellis, 2003; Netemeyer et al., 2003).

Temporal stability, also known as test-retest reliability, was assessed by Pearson's correlation coefficients between the same scale measured for the same population at two different points in time (DeVellis, 2003). In our case, we calculated these correlations for each consecutive survey wave, as well as between the 1991 and 2006 surveys.

Criterion-related validity is typically assessed by comparing the scale's degree of agreement with a "gold standard" measurement (DeVellis, 2003). In the case of measuring urbanicity, there was no accepted gold standard so, for criterion-related validity, we compared our scale to the existing official classification of urban status using a dichotomous classification as rural or urban, as well as a four category classification of urban city, county township, suburb, and village. It is expected that these correlations will be attenuated compared expected results from comparison to a better "gold standard" measure (Willett, 1998). To compare our measure to

the dichotomous urban/rural classifications, we dichotomized our measure based on the receiver operator characteristic curve determination of maximized sensitivity and specificity for the year 2000 and compared the resulting categories using the kappa statistic for agreement beyond chance. We assessed criterion-related validity using Spearman's rank correlation coefficient and the four-category measure mentioned above.

As a sensitivity test of the scale to our *a priori* choice of equally weighted scoring of the components, we compared the results of the scores based on *a priori* scoring methods (equally weighted components) to the scale derived from data-driven factor analysis techniques (components weighted according to factor analysis weights). We performed a factor analysis of the 12 components and used the factor weights as multipliers for each component so that the total score reflected unequal influence of the various components. We then performed the equivalent criterion-related validity and temporal stability testing for the scores from factor-analysis-derived weights.

The construct validity, the extent to which the scale coincided with the phenomena known to differ by urban status (DeVellis, 2003), was assessed by regression of urbanicity on per capita inflation-adjusted household income, as well as on odds of having more than one child. Household income is consistently related to urbanicity in China, making it a suitable regressor for an evaluation of construct validity. China's one-child policy is well known; however, in many provinces there exists a "1.5 child policy", whereby only rural couples are allowed to have a second child if their first child is a girl. This means that in rural places we would expect a higher average number of children, making this also a good regressor for the evaluation of the construct validity. We performed the regression using quintiles categories of the scale, modeled as indicators variables.

To examine the additional uses of the new continuous scale that would not be possible with the static urban/rural dichotomous variable, we looked descriptively at the heterogeneity in the changes among the scale components. To do this, we first grouped communities by their initial score on the urbanicity scale and then further group them by the level of change in urbanicity score over the survey years. We calculated average change in each of the urbanicity scale components for each category of baseline and change in urbanicity score. We also compared estimates of incident overweight/obesity status by the initial level of urbanicity according to the new scale and by change in the level of urbanicity using logistic regression (with robust standard errors) for a subset of women who had BMI measured in both 1991 and 2004 and who were not overweight/obese in 1991 (n=1433). Fifty percent of the women who were eligible in 1991 (n=2861) were followed up in 2004. The loss to follow-up occurs because CHNS does not follow families or individuals who move out of the community. The women who left the sample did not have significantly different BMI at baseline, but they were significantly younger at baseline and were more like to reside in an urban community; these characteristics are included in the model. The model used the urbanicity scale in its continuous form and the squared term for this variable to allow nonlinearity. It also included change in urbanicity and its square. It included a number of interaction terms to allow heterogeneity of effects (baseline urbanicity by urbanicity change, baseline urbanicity squared by urbanicity change, urbanicity change squared by baseline urbanicity, and baseline urbanicity squared by urbanicity change squared). For comparison, we also used logistic regression to estimate the incident overweight/obesity value in the same sample but using only the static urban/rural dichotomous variable.

As a final example of an additional use of the new continuous scale, we conducted an individual-level fixed-effects regression to examine the change in the percentage of the total calories from fat in relation to the change in community urbanicity. Fixed-effects regression compares each individual to herself over time and in so doing controls for individual-level,

time-invariant confounders (Angrist & Jorn-Steffen, 2008). Fixed-effects regression also requires that the independent variables in the model change over time for at least some of the population. The new urban scale satisfied this criterion, whereas the comparison measure, the urban/rural dichotomous variable, was static and could not be used in such a regression. The fixed-effect model regressed the percentage of calories from fat on the urbanicity level of each individual's community. We included the squared term for urbanicity.

## RESULTS

The number of communities surveyed and calendar year for each wave are displayed in Table 1a. The 12 components of the scale appear to represent a unidimensional underlying construct, which we called urbanicity, as evidenced by the extremely high eigenvalue of only one factor in the exploratory factor analysis of the components. This result is reflected in the scree plot for the factor analysis displayed in Figure 1. The mean scores on the scale by year and further stratified by community type are displayed in Figure 2a. The distributions on the scale for year 2006 are given in Figure 2b. The supplementary material provides means and standard deviations for each of the components by year and by official classification of city, town, suburb, or village [LINK TO SUPPLEMENTARY MATERIAL FILE].

Cronbach's alpha is very high for all years of the survey ( $\alpha=0.85$  to  $0.89$ ), indicating the scale had very good internal consistency (Table 1a) (Netemeyer et al., 2003). The corrected item-scale correlations, which calculated the correlation of each component with the remaining components in the scale, were also generally quite good, with correlations  $>0.40$  to  $0.50$  (Table 1a).

Moreover, the scale exhibited temporal stability in the test-retest reliability scenarios (Table 1b). Correlation between each of the consecutive waves was quite high ( $r=0.90$  to  $0.94$ ). Additionally, we had expected the correlation between two more distant time points to be somewhat less than the correlations between two more proximal time points; the correlation between 1991 and 2006 scores was  $0.84$ .

There is evidence for the scale's criterion-related validity from its comparison to the official classification of communities as urban or rural. The kappa statistic for agreement of our scale and the "gold standard" ranged from fair to good interscale agreement (Table 1b). Spearman's rank correlation also supported criterion-related validity (Table 1b).

The criterion-related validity and the temporal stability were largely unchanged when we substituted the factor-score weighting of the components for equally weighted components in our sensitivity analysis (results not shown).

The mean urban scores for communities increased over time, as expected (Figure 2). The linear and logistic regressions used to assess construct validity indicated that our scale demonstrated good properties on this dimension as well (Tables 2 and 3). Increasing scores on the urban scale were significantly associated with increases in the adjusted per capita household income and with a significantly lower odds of having more than one child.

Figure 3 displays the heterogeneity of change in urbanicity components from 1991 to 2006. Overall change in score was calculated for communities that were included in the survey in both 1991 and 2006 ( $N=176$ ). Communities were classified by their baseline urban score in 1991, using a score of 50 as a cut-off. Figure 3a represents the changes in urban score for communities that had lower initial scores in 1991 ( $N=106$ , 60%). Figure 3b presents the same information for communities with higher initial urban scores ( $N=70$ , 40%). Communities were grouped by overall level of change in urban score, in categories of a net decrease in urban score (urban score change  $<0$ ), or an increase in score of 0 to 15, 15 to 30 points, and  $>30$  points.



Communities with higher scores on the urban scale in 1991 (>50 points) had a slightly higher overall mean change in score of this time period compared to those classified as urban (urban mean change=21, rural mean change=16,  $p$  value<0.08). Furthermore, a larger percentage of the more urban communities fell into the two highest categories of change (74% versus 54%), and a similar proportion of the urban communities fell into the lowest category of change (4.3% versus 3.8%). Of note, when evaluating growth or declines in the separate components we saw that, across community-type and across levels of urban score change, health infrastructure (health care facilities and pharmacies) appears to have consistently declined or experienced only very small growth. However, sanitation and housing infrastructure consistently experienced fairly large gains across all levels of initial and change in urban scores.

Among the more rural communities that experienced a decrease in urban score or a smaller level of increase, it appears that declines in scores were largely reflective of loss of both modern and traditional markets in these places, as well as losses in health infrastructure. Similarly, among the more urban communities, for those that experienced a net loss in urban score, the losses seemed to be driven by these same components, plus additional losses in the transportation component.

Figure 4 displays the odds ratios of incident overweight/obesity over the period 1991 to 2004 for the subset of the female population who had BMI measured during both of these years ( $n=1433$ ) (Figure 4). When comparing the odds ratio for overweight/obesity using only the traditional dichotomous classification, the odds of overweight/obesity for urban women were not statistically different from the odds for rural women (OR 1.13 (0.78, 1.66)). The variation that can be seen by using the baseline and change in urban score from the new scale reveals how the new scale added value over the traditional categorization. The odds of overweight are significantly higher for women in communities with greater baseline and change in urban score compared to those in the least urban communities that did not change in urbanicity. Details of these comparisons are elaborated in the discussion.

The fixed-effects regression of change in a community's level of urbanicity and change in an individual's percentage of total energy from fat explored more proximate dietary changes in relation to urban places. This regression indicated that increasing community urbanicity is associated with increasing energy from fat ( $\beta=0.27$  (0.21, 0.32)), but that the effect is smaller with larger changes in urbanicity ( $\beta_{\text{urbanicity squared}}=-0.0011$  (-0.0015, -0.0065)).

## DISCUSSION

Major social, demographic and economic change has occurred in both urban and rural areas across the developing world. Since 1980, China has recorded both some of the most rapid economic changes, along with vast migration and urbanization (Ma, 2002). This paper has developed and tested the reliability and validity of a new scale to measure urbanicity on a continuum in China. We demonstrated unidimensionality of the scale, which is desirable since we aimed to tap the latent construct of urbanicity. The scale performed very well on tests of reliability, content, criterion, and construct validity. Additionally, the new scale added valuable information to the analyses of differences in health across levels of urbanicity.

Van de Poel and colleagues (2009) used factor analysis to create a similar tool to measure urbanicity using the CHNS data. Many of the scale development and testing procedures that we presented were not provided in the Van de Poel et al. (2009) analysis (content validity, Cronbach's alpha, item-scale correlations, test-retest reliability), so direct comparison on these aspects was not possible. Both scales used household income for a test of construct validity and tracked the change in urban score over time as an assessment of temporal stability and the scales performed similarly. However, although the Van de Poel urban index performed well

on a number of criteria, a noted limitation of that index is that the average scores for cities and towns are the same in the most recent survey years. By comparison, the average scores for cities and towns were significantly different using our scale (Figure 2), which reflects our attempts to incorporate modern features that might only be available in the most urban places.

Our new scale addressed urbanicity on a continuum thereby allowing us to assess the degree of change in urban features across communities with varying baseline and change in urban features. Our descriptive findings presented in Figures 3a and 3b display the heterogeneities and the similarities in dynamic urban features across different types of communities. The analysis of incident overweight/obesity status briefly exemplifies the heterogeneity that may underlie relationships between baseline and change in urbanicity of a place and health. The analysis is conducted among women who were not overweight/obese in 1991 (and thus at risk for becoming overweight/obese) and were followed up in 2004. The results examine expected odds of becoming overweight/obese for women depending on their community's initial degree of urbanicity and its level of change in urbanicity over the follow-up period.

This analysis finds that, for women living in communities with a low initial urban score (~30), a higher level of change in urban score over the follow-up years (vs no change) is associated with a statistically significantly higher odds of incident overweight/obesity during that time (*p-value* for 10, 20, 30 point increase vs 0 point increase= 0.03 for each). Women living in communities in the midrange of urban scores (~50) experience odds of overweight/obesity that are not statistically significantly different from the women in communities with initially lower urbanicity (30–40) that experienced a 20–30 point increase in urbanicity over the study period. Increasing urbanicity among this group does not confer significantly greater odds of overweight/obesity. Additionally, for women in communities with the highest initial urban scores (~60 to 70) and no change in urban score, the odds of incident overweight/obesity are still statistically significantly higher than for less urban communities (initial scores of 30), but a higher change in urban score (increase of 10–15 points) among this group also does not predict statistically significantly higher odds of incident overweight/obesity. It appears that increased urbanicity is associated with increased overweight/obesity up to a certain point, but increases in urbanicity beyond that point do not confer greater odds of incident overweight/obesity.

This relation certainly deserves further exploration and for now, we speculate a few possible explanations. Perhaps the urban features most associated with changes in weight and risk for overweight/obesity, such as occupational changes and changes to traditional diets, are some of the first manifestations of urbanicity. In highly urbanized environments, it could be that social norms get established and relative adult overweight/obesity risk declines-- perhaps Western media influences and body ideal are stronger, or have persisted longer, in the most urban places and act to decrease overweight/obesity incidence in the most urbanized places.

As noted in the methods section, there is potential for selection bias as those lost might have behaved differentially despite beginning with the same baseline BMI ( $p>0.05$ ). In a sensitivity analysis, we ran a probit Heckman selection model and found that accounting for potential selection bias did not change the sign or significance of the coefficients in the model (Hausman, 1978; Heckman, 1979).

The analysis of overweight/obesity and urbanicity is one example of the utility of examining urbanicity on a continuum in relation to health outcomes. By comparison, the results from the analogous logistic regression, which included the static urban-rural dichotomous variable, indicated that there was no significant difference between the odds of overweight/obesity for urban places compared to rural places and, as such, misses the heterogeneity that may be operating in places of varying degrees of urbanicity. Finally, the new urban scale is time varying and thus can be used in change analyses, such as the econometric fixed-effects regression of

percent energy of fat changing as a function of change in community urbanicity. The fixed-effects regression also indicated that change in community urbanicity is associated with increases in fat in the diet but that this effect diminishes with the largest changes in urbanicity.

The scale has many strengths but limitations should also be noted. First, because of some missing values for communities in each year, we had to impute a small portion of the data points. Since the largest number of imputations occurred for the first wave in 1989, we repeated the tests of reliability and validity excluding this wave, and none of the results changed substantially. We also repeated these analyses excluding the component “traditional markets.” The reliability was essentially unchanged when traditional markets was excluded, but the criterion-related validity was slightly lower with exclusion of these markets, so we concluded that their inclusion was beneficial to the overall scale.

As mentioned previously, a few variables were not measured for all years. These variables represent recent modern additions to communities, such as new food vendor and supermarket types, cell phones, and computers. We deemed these important contributions to the scale and therefore included them, but small complications arise in the comparability of the scale across years. We have attempted to make the scores comparable across years, as described below.

To account for the fact that some of the food vendors in the “modern market” component were measured beginning in only 2004, we adjusted the potential possible points in this component so that in years prior to 2004 the collected markets could give a total possible 10 points, while in years after 2004, these markets were weighted down so that new markets could be included and the total possible would still be 10 points.

For the “social services” component, which included day care and community health insurance availability, and which we collected beginning in 2000, as well as for computers and cell phones, which we began querying in 1997 and 2004, respectively, we conceptualized the true levels of these as being close to nonexistent before we began collecting the data. Therefore, we kept the total possible points consistent across years.

An additional limitation is that the construct of urbanicity may overlap with other important constructs, such as economic development. In our scale development, we conceptualized urbanicity as the latent construct that produces these measurable conditions that act as indicators of this latent construct. We built our scale from descriptions of urban features in the literature in attempt to capture this construct in particular. Because of the overlap in definitions, we cannot assuredly separate these two constructs; however, if two semi-distinct constructs were at work, we would have expected the test of dimensionality to reveal a multidimensional scale rather than a unidimensional scale. The finding of unidimensionality helps assuage doubts that we may be tapping two separate constructs. Building the scale around the defining and distinguishing features of urban places gives confidence that the one construct is in fact urbanicity.

The scale developed in this paper has the potential for use in many analyses beyond this brief example. It could be used to further probe other health outcomes, consumption behaviors, and demographic and economic activities, to name a few. After this paper is published, the scale and full documentation will be publically available to researchers using the public-use CHNS data. Additionally, as previously mentioned, using the new urban scale as a covariate will enable less misclassification measurement error compared to the dichotomous classification, even when urbanicity is not the explanatory variable of interest.

In summary, the urbanicity scale developed and validated for CHNS captures enormous changes in context across time and space. It performed very well on all conducted tests of internal consistency, reliability, content validity, criterion-related validity, and construct

validity. The overall scale promises to be most useful for other economic, demographic, social welfare, and health outcomes.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

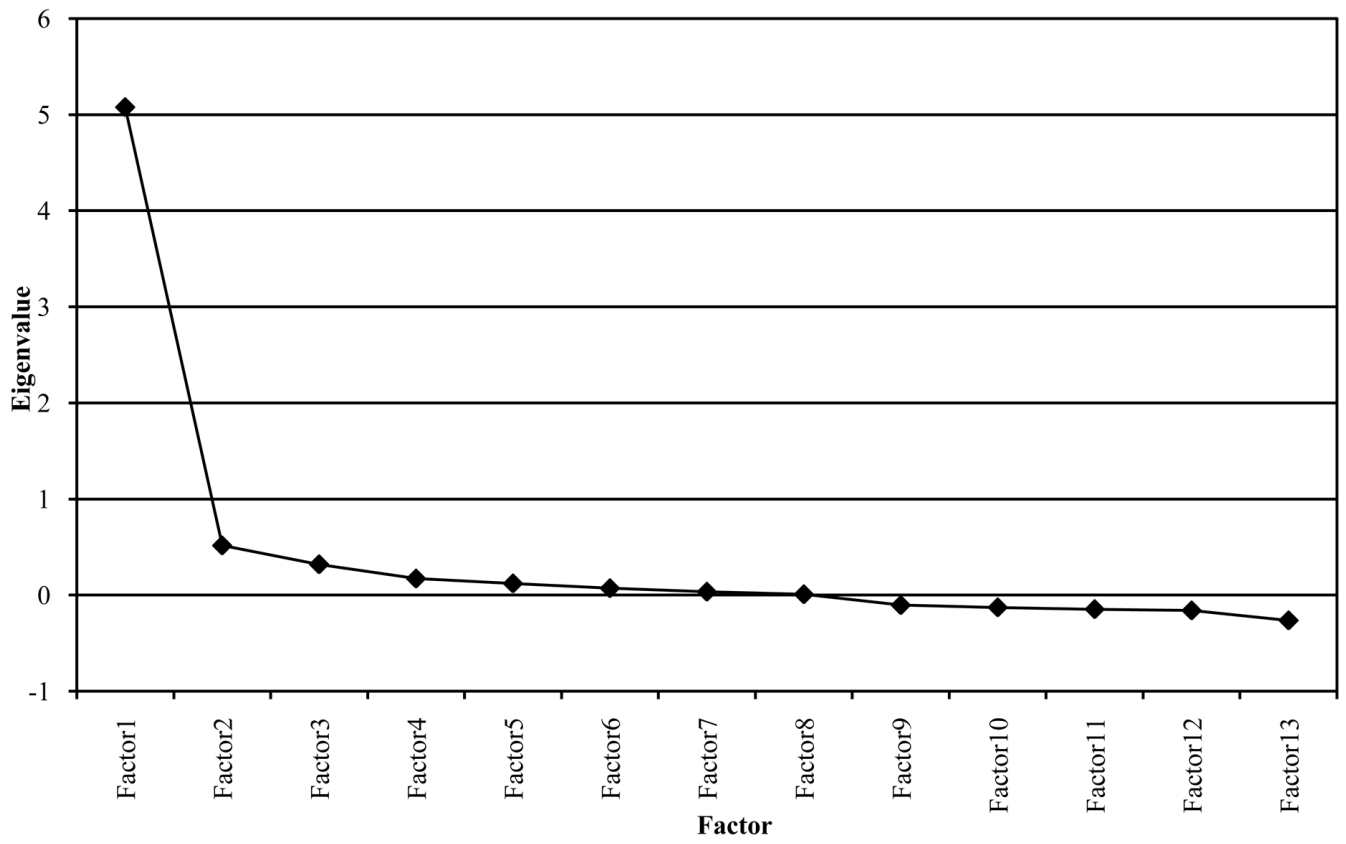
## Acknowledgments

This research uses data from China Health and Nutrition Survey (CHNS). We thank the National Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center, the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, and R01-HD38700) and the Fogarty International Center, NIH for financial support for the CHNS data collection and analysis files from 1989 to 2006. Intellectual conceptualization of this urbanicity measure began with Jodi Stookey and later Michelle Mendez, and we thank both for earlier work. We also wish to thank Jim Terry for the tremendous programming support, without which this project would not have been possible, Shufa Du for help with understanding many China-related issues, Tom Swasey for graphics support and Frances Dancy for administrative support.

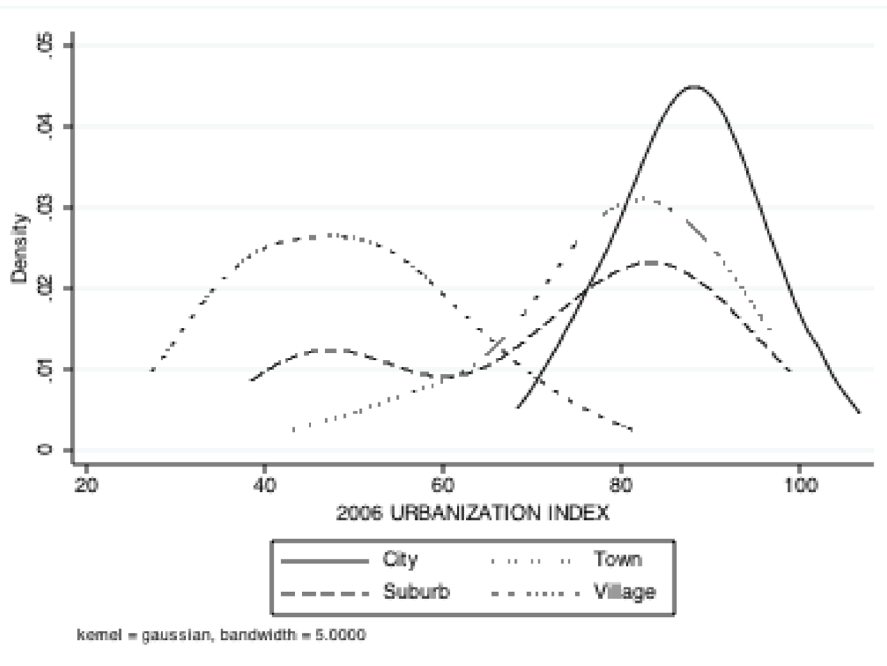
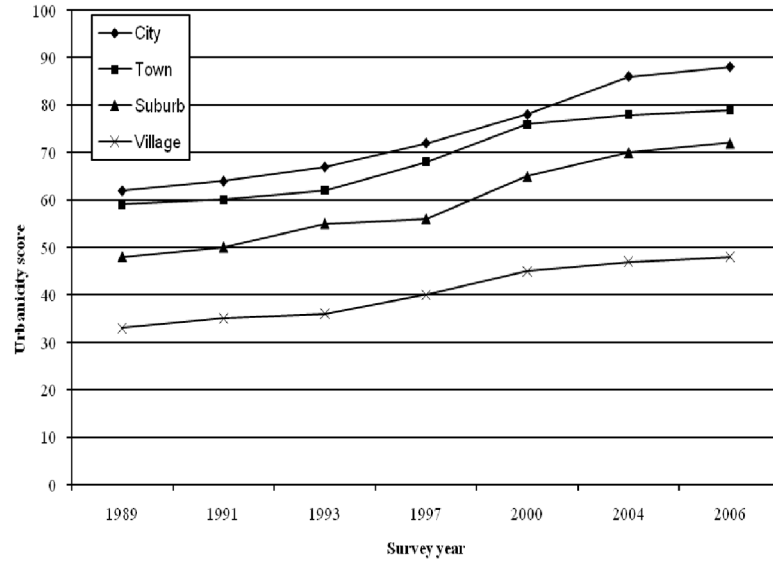
## References

- Allender S, Foster C, Hutchinson L, Arambepola C. Quantification of Urbanization in Relation to Chronic Diseases in Developing Countries: A Systematic Review. *Journal of Urban Health* 2008;85(6):938–951. [PubMed: 18931915]
- Angrist, JD.; Jorn-Steffen, P. *Mostly Harmless Econometrics: An Empirist's Companion*. Princeton: Princeton University Press; 2008.
- Batal M, Boulghourjian C, Abdallah A, Afifi R. Breast-feeding and feeding practices of infants in a developing country: a national survey in Lebanon. *Public Health Nutr* 2006;9(3):313–319. [PubMed: 16684382]
- Brockhoff M, Brennan E. The Poverty of Cities in Developing Regions. *Population and Development Review* 1998;24(1):75–114.
- Champion, AG.; Hugo, G., editors. *New Forms of Urbanization: Beyond the Urban-Rural Dichotomy*. Aldershot, England: Ashgate Publishing Limited; 2004.
- Dahly DL, Adair LS. Quantifying the urban environment: a scale measure of urbanicity outperforms the urban-rural dichotomy. *Social Science and Medicine* 2007;64(7):1407–1419. [PubMed: 17196724]
- DeVellis, R. *Scale Development: Theory and Applications*. Thousand Oaks, CA: Sage Publications; 2003.
- Fotso JC. Child health inequities in developing countries: differences across urban and rural areas. *Int J Equity Health* 2006;5:9. [PubMed: 16831231]
- Friedman, J. *China's Urban Transition*. Minneapolis, MN: University of Minnesota Press; 2005.
- Hausman JA. Specification Tests in Econometrics. *Econometrica* 1978;46(6):1251–1271.
- Heckman JJ. Sample Selection Bias as a Specification Error. *Econometrica* 1979;47(1):153–161.
- Hinde PR, Mturi AJ. Social and economic factors related to breast-feeding durations in Tanzania. *J Biosoc Sci* 1996;28(3):347–354. [PubMed: 8698714]
- Hu X, Cook S, Salazar MA. Internal migration and health in China. *Lancet* 2008;372(9651):1717–1719. [PubMed: 18930533]
- Johansson SR, Mosk C. Exposure, Resistance and Life Expectancy: Disease and Death during the Economic Development of Japan, 1900–1960. *Population Studies* 1987;41(2):207–235. [PubMed: 11621337]
- Ma LJC. Urban transformation in China, 1949–2000: a review and research agenda. *Environment and Planning A* 2002;34:1545–1569.
- McDade TW, Adair LS. Defining the “urban” in urbanization and health: a factor analysis approach. *Social Science & Medicine* 2001;53(1):55–70. [PubMed: 11380161]
- Mendez MA, Monteiro CA, Popkin BM. Overweight exceeds underweight among women in most developing countries. *American Journal of Clinical Nutrition* 2005;81(3):714–721. [PubMed: 15755843]

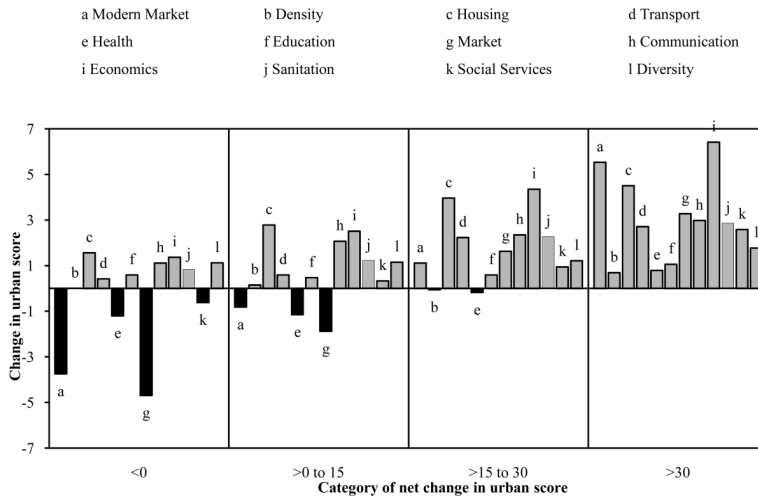
- Mendez, MA.; Popkin, B. Globalization of food systems in developing countries: impact on food security and nutrition. Rome, Italy: Food and Agriculture Organization, Food and Nutrition Division; 2005. Globalization, urbanization, and nutritional change in the developing world; p. 55-80.
- Monda KL, Gordon-Larsen P, Stevens J, Popkin BM. China's transition: the effect of rapid urbanization on adult occupational physical activity. *Soc Sci Med* 2007;64(4):858–870. [PubMed: 17125897]
- Montgomery, MR.; Ezeh, AC. The Health of Urban Populations in Developing Countries: An Overview. In: Galea, S.; Vlahov, D., editors. *Handbook of urban health: Populations, methods, and practice*. New York: Springer; 2005.
- National Research Council. *Cities Transformed: Demographic Change and Its Implications in the Developing World*. In: Montgomery, MR.; Stern, R.; Cohen, B.; Reed, HE., editors. *Panel on Urban Population Dynamics*. Washington, DC: Committee on Population, Division of Behavioral and Social Sciences and Education; 2003.
- Netemeyer, R.; Bearden, W.; Sharma, S. *Scaling Procedures: Issues and Applications*. Sage Publications; 2003.
- Popkin BM, Bilsborrow RE, Akin JS. Breast-feeding patterns in low-income countries. *Science* 1982;218(4577):1088–1093. [PubMed: 7146896]
- Popkin BM, Du S, Zhai F, Zhang B. Cohort Profile: The China Health and Nutrition Survey--monitoring and understanding socio-economic and health change in China, 1989–2011. *Int J Epidemiol* 2009;dyp322.
- Poston DL, Mao MX. Interprovincial migration in China, 1985–1990. *Res Rural Sociol Dev* 1998;7:227–250. [PubMed: 12294798]
- Preston SH, Walle Evd. Urban French Mortality in the Nineteenth Century. *Population Studies* 1978;32(2):275–297. [PubMed: 11630579]
- Rondinelli, DA. *Secondary Cities in Developing Countries: Policies for Diffusing Urbanization*. Beverly Hills, CA: Sage Publications; 1983.
- Rondinelli DA. Cities as agricultural markets. *The Geographical Review* 1987;77(4):408–420.
- Shen J. Rural development and rural to urban migration in China 1978–1990. *Geoforum* 1995;26(4):395–409. [PubMed: 12348193]
- Van de Poel E, O'Donnell O, Van Doorslaer E. Are urban children really healthier? Evidence from 47 developing countries. *Social Science & Medicine* 2007;65(10):1986–2003. [PubMed: 17698272]
- Van de Poel E, O'Donnell O, Van Doorslaer E. Urbanization and the spread of diseases of affluence in China. *Economics & Human Biology* 2009;7(2):200–216. [PubMed: 19560989]
- Vlahov D, Galea S. Urbanization, urbanicity, and health. *J Urban Health* 2002;79(4 Suppl 1):S1–S12. [PubMed: 12473694]
- WHO/FAO. *Expert Consultation on Diet, Nutrition and the Prevention of Chronic Diseases Report of the joint WHO/FAO expert consultation*. Geneva: World Health Organization; 2003.
- Willett, W. *Nutrition Epidemiology Second Edition*. New York: Oxford University Press; 1998.
- Yang X, Guo F. Gender differences in determinants of temporary labor migration in China: a multilevel analysis. *Int Migr Rev* 1999;33(4):929–953. [PubMed: 12349706]
- Zhu, Y. *Changing Urbanization Processes and In Situ Rural-Urban Transformation: Reflections on China's Settlement Definitions*. In: Champion, AG.; Hugo, G., editors. *New Forms of Urbanization: Beyond the Urban-Rural Dichotomy*. Aldershot, England: Ashgate Publishing Limited; 2004.



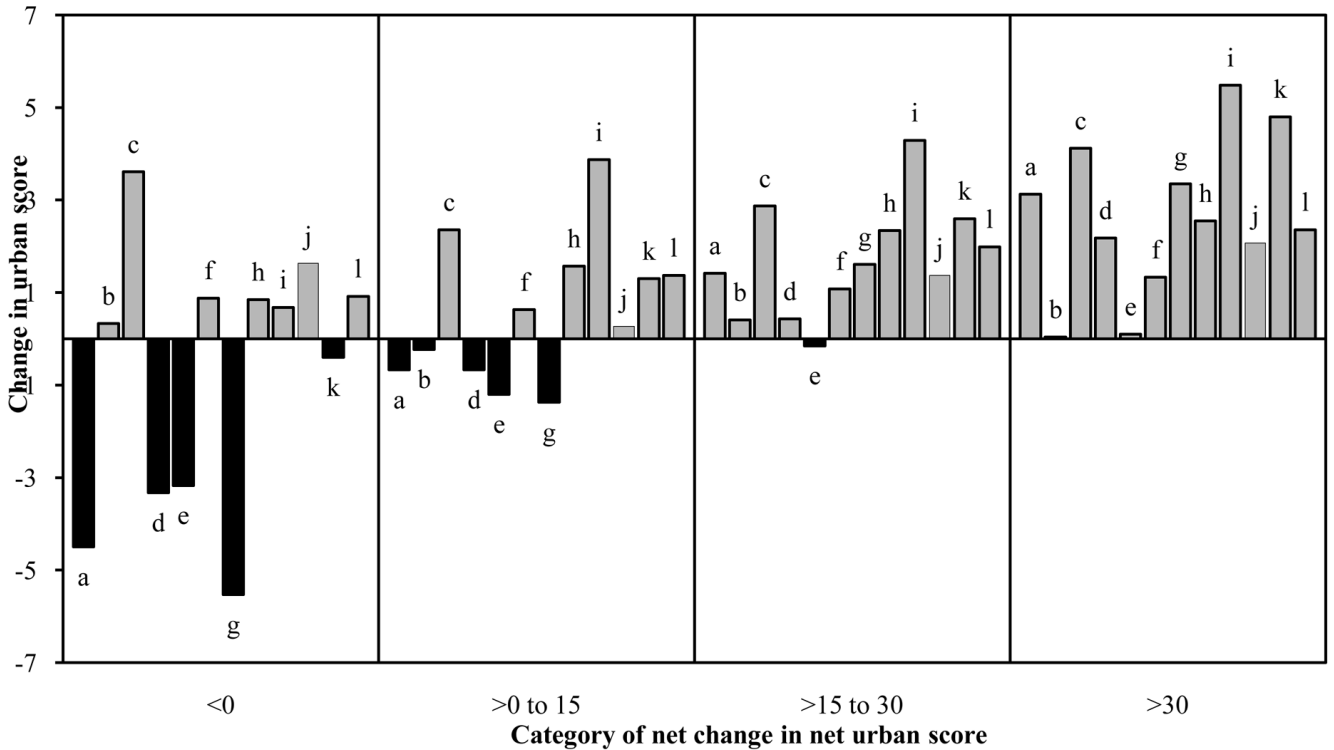
**Figure 1.** Scree plot of eigenvalues after factor analysis of 12 urban components.



**Figure 2.**  
 Figure 2a. Mean urbanicity scores by community type: CHNS 1989 – 2006.  
 Figure 2b. Distribution of urbanicity scores in 2006 by official designation (city, town, suburb, village).



a Modern Market                      b Density                      c Housing                      d Transport  
 e Health                                  f Education                      g Market                      h Communication  
 i Economics                              j Sanitation                      k Social Services              l Diversity

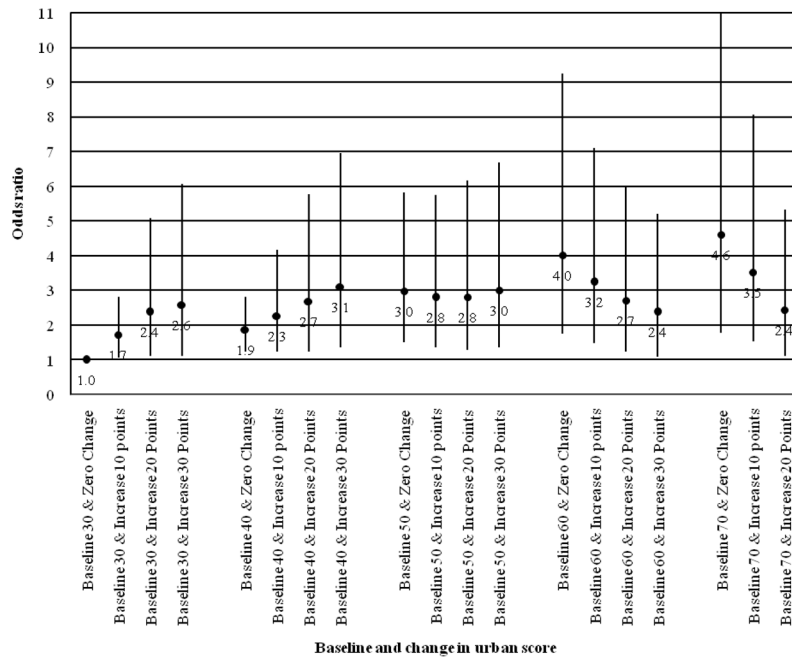


**Figure 3.**

Figure 3a. Heterogeneity in urban change for “rural” communities (initial urbanicity score in 1991  $\leq$  50): change in individual components by level of overall change in urban score (1991–2006).

Figure 3b. Heterogeneity in urban change for “urban” communities (initial urbanicity score in 1991  $>$  50): change in individual components by level of overall change in urban score (1991–2006).





**Figure 4.** Odds ratios and 95% confidence intervals of incident overweight/obesity for females by community’s baseline and net change in urbanicity score 1991–2004.

Table 1a

Urbanicity scale properties by year: Cronbach's alpha, item-scale correlations

Year	Communities (N)	Alpha	Internal consistency: Corrected item-scale correlations by survey year											
			Population Density	Education	Sanitation	Housing	Transportation	Communications	Health Infrastructure	Traditional Markets	Economy	Diversity	Modern Markets	Social Service
1989	190	0.86	0.45	0.59	0.70	0.68	0.62	0.66	0.61	0.52	0.66	0.60	0.40	NA
1991	190	0.85	0.43	0.61	0.71	0.72	0.46	0.64	0.68	0.52	0.73	0.60	0.43	NA
1993	188	0.86	0.40	0.63	0.73	0.75	0.59	0.70	0.63	0.45	0.65	0.57	0.48	NA
1997	192	0.87	0.37	0.68	0.75	0.78	0.59	0.68	0.61	0.50	0.64	0.65	0.58	NA
2000	217	0.88	0.39	0.65	0.75	0.81	0.49	0.65	0.50	0.44	0.68	0.6	0.58	0.31
2004	216	0.89	0.49	0.65	0.76	0.80	0.39	0.65	0.61	0.55	0.68	0.71	0.72	0.59
2006	218	0.89	0.47	0.67	0.77	0.80	0.40	0.68	0.62	0.61	0.7	0.67	0.75	0.51

**Table 1b**

Urbanicity scale properties by year: criterion validity, and test-retest reliability

Year	Criterion validity: Kappa statistic for agreement beyond chance			Kappa Statistic	Spearman's Correlation	Temporal Stability: Test-retest reliability	
	Expected Agreement	Observed Agreement				Comparison Years	Pearson's Correlation
1989	63%	71%		0.21	0.78		
1991	62%	72%		0.25	0.77	1989–1991	0.94
1993	60%	75%		0.38	0.78	1991–1993	0.91
1997	56%	69%		0.29	0.75	1993–1997	0.90
2000	52%	74%		0.45	0.78	1997–2000	0.90
2004	51%	77%		0.53	0.77	2000–2004	0.92
2006	51%	74%		0.48	0.76	2004–2006	0.93

All correlations are significant at 5% level

**Table 2**

Cross-sectional linear regression of urban scale quintiles on adjusted household per capita

Quintile	Coefficient (Robust SE)		
	1991(n=3607)	2000(n=4313)	2006(n=4302)
Urban Scale Quintile 1(referent)	NA	NA	NA
Urban Scale Quintile 2	709** (177)	837** (311)	469 (625)
Urban Scale Quintile 3	1021** (201)	1670** (334)	1235* (631)
Urban Scale Quintile 4	1424** (206)	2749** (391)	5090** (1099)
Urban Scale Quintile 5	1957** (175)	3625** (443)	5560** (917)

\*  
 $p$  value  $\leq 0.05$ \*\*  
 $p$  value  $\leq 0.01$

**Table 3**

Cross-sectional logistic regression of urban scale quintiles on odds of having more than one child

Quintile	Odds Ratio (Robust SE)		
	1993 (n=2565)	2004 (n=2752)	2006 (n=2636)
Urban Scale Quintile 1 (referent)	NA	NA	NA
Urban Scale Quintile 2	0.73 (0.20)	0.40** (0.11)	1.12 (0.21)
Urban Scale Quintile 3	0.52** (0.14)	0.53** (0.11)	0.79 (0.16)
Urban Scale Quintile 4	0.38** (0.10)	0.23** (0.05)	0.26** (0.07)
Urban Scale Quintile 5	0.11** (0.02)	0.08** (0.02)	0.13** (0.04)

\*\*  
p value<0.01