Economic Development, Urbanicity and Adult Health Changes in China

Jessica Jones-Smith and Barry M Popkin

Corresponding author:
Jesse Jones-Smith
Carolina Population Center
CB # 8120 University Square
University of North Carolina at Chapel Hill
Chapel Hill, NC 27516-3997
Phone: (919) 966-1499 or (415)722-3978
Fax: 919-966-9159
E-MAIL: jessejonessmith@unc.edu

Introduction

Rapid economic development in lower income countries has led to vast changes in physical, social, demographic and economic landscapes. Subsequently, these changes are thought to have impacts on the social welfare and health of hundreds of millions of individuals. People have migrated to big and small cities, roads have been paved, electricity delivered. More people are finding employment outside of agriculture; public institutions, such as schools and hospitals, have increased in number, and commercial markets have also increased in number and diversity(United Nations 2002; National Research Council 2003). Many of these manifested changes could be considered components of the underlying construct of urbanization and urbanicity.

Urbanization has been defined as the change in size, density and heterogeneity in places and the migration into cities, while urbanicity is descirbed as the degree to which a place exhibits urban characteristics (Vlahov and Galea 2002). The distinguishing features of urban places have been described in the literature, and include: population and proximity, infrastructure, diversity, social environment, culture and economic activity(Vlahov and Galea 2002; National Research Council 2003). These definitions both invoke movement, change and fluidity. However, most working definitions of urban and rural are based on an absolute threshold of population and population density, such as >50,000 people and density of >1,000 people per square mile. Clearly, these definitions could not distinguish cities on a continuum of urbanization or urbanicity (Brockerhoff 1998; McDade and Adair 2001; Mendez and Popkin 2005).

Nevertheless, crude classification of geographic areas as urban or rural has revealed important differences in health based on urban/rural status (Brockerhoff 1998; Vlahov and Galea 2002; Mendez, Monteiro et al. 2005; Fotso 2006; Van de Poel, O'Donnell et al. 2007) For example, infant mortality, obesity, breast-feeding studies have all found urban residence to represent a critical dimension (Popkin, Bilsborrow et al. 1982; Brockerhoff 1995; Hinde and Mturi 1996; Mendez, Monteiro et al. 2005; Batal, Boulghourjian et al. 2006). 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 and may require deeper examination (McDade and Adair 2001). Trajectories of development likely differ spatially and temporally and by degree of baseline urbanicity. Many villages throughout the developing world are gaining factories or only selected dimensions of what we might say constitute an urban environment, while other demographic entitites are losing factories, but gaining universities, markets, transportation infrastructure (Rondinelli 1983; Rondinelli 1987). Economic, social and demographic change is complex and does not proceed in a monotonic shift forwards or backwards.

As theories continue to develop to explain the effects of environments-both urban and rural- on health, there is a need for methods to provide a more refined measurement of urbanicity and its components. Additionally, a refined measure of urbanicity would enable both inter- and intra- urban and rural comparisons, which is a relatively untapped area of research for demographers(National Research Council 2003). An improved measure of urbanicity would draw information from the domains that characterize urban versus rural places, would be sensitive to changes over time and would represent gradations on the continuum from rural to urban environments.

In this paper, we develop a multi-component scale to measure urban features on a continuum. We evaluate the new scale for reliability, and content, criterion-related and predictive validity. Finally, we provide brief examples to demonstrate that the scale provides a more nuanced and meaningful understanding of the ways urbanicity may change over time and how this may, in turn, affect adult health outcomes. We use incident overweight/obesity as an adult health outcome, which can be particularly useful as it captures both the effects of economic, mass media and food market shifts on food selection and the effects of transportation, technology and other societal changes on energy expenditure.

China provides a unique of opportunity to develop and test such a scale. Extensive community-level data have been collect from 1989 to 2006 in the China Health and Nutrition Survey (CHNS). Furthermore, China has experienced urbanization relatively recently with marked migration to bigger towns and cities beginning in the early to mid-1980's. Of course, this is not the first period of migration—forced or unforced—in China's history as the cultural revolution saw millions shifted from cities to rural areas but these changes were caused by political rather than social and economic changes. The CHNS captures the impact of China's reforms, including: elimination of most food subsidies and reduction of remaining food subsidies along with those for fuel and most other basic commodities by 1991; freely fluctuating market prices; privatization of many state enterprises; permitting and creating the conditions for the establishment of a private sector in most areas of economic activity; and initiation of 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 subsequently to cities and towns across the country (Poston and Mao 1998; Yang and Guo 1999; Hu, Cook et al. 2008).

Methods

The China Health and Nutrition Survey data collection began in 1989, and has been implemented every 2-4 years since. The CHNS uses a multistage cluster sample design to survey individuals and households within 218 neighborhoods within 9 provinces in China. These 9 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 to form 3 income groups—low, medium, high. A weighted sample of 4 counties was selected, as was the provincial capital city and a lower income city in each province (in some provinces, a large city other than the provincial capital had to be substituted). Within these cities or counties, neighborhoods were randomly selected, resulting, originally, in 190 communities, and, currently, in 218 communities. Within the communities, households were selected randomly from a carefully collected household roster and all members in each household were interviewed. The household roster was used to follow-up each of the originally sampled households for subsequent survey panels. If a member of a respondent household formed a new household in the same neighborhood, this household was added to the sample. In addition to individual-level data, detailed information on community infrastructure, services, demographic and economic environment are collected. These rich data enable the development of a scale based on many components that are theorized to capture the underlying latent construct of urbanicity.[i]

The CHNS includes extensive individual, household and community level surveys; the current analysis utilizes information from each of these surveys. The household survey was conducted inside the participant's home by trained field workers over 3-4 days; the individual survey was conducted during this time as well. The information for community survey was collected by trained field workers, during which the interviewer obtained information from a combination of neighborhood officials, informants, visits to markets, and official records.

<u>Scale Construction</u>

Variable Selection
In order to determine which variables should be used to construct the urbanicity scale, we consulted existing 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 (Vlahov and Galea 2002; National Research Council 2003). From these sources we compiled a list of characteristics that are thought to distinguish urban places from rural places. This includes: 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); 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. We identified thirteen broad areas, which we call components, to distinguish urbanicity. Multiple variables were used to construct some of these components, while 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 seems to capture the content domain of urbanicity (DeVellis 2003). The evaluators of content domain 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 thirteen separate components thought to define and distinguish urbanicity that could be operationalized in the CHNS data, which we name: Population; Population Density; Economic Activity; Traditional Markets; Modern Markets; Transportation Infrastructure; Sanitation; Communications; Housing; Education; Diversity; Health Infrastructure; Social Services. The variables included in each component are as follows:

- Population: Total population of the community from official records.
- Population Density: Total Population of the community divided by community area, from official records.
- Economic Activity: Ordinary wage for male worker and percent of the population engaged in nonagricultural work.
- Traditional Markets: Availability of, distance from and days of operation for 8 different types of market (including various 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.
- Transportation Infrastructure: Most common type of road, presence of and distance to bus stop, and presence of and distance to train stop.
- Sanitation: Prevalence of households with treated water, and prevalence of households without excreta present outside the home.

- Communications: Availability in community 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, percent of community that cooks with gas.
- Education: Average education 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 the community, number of pharmacies in community.
- Social Services: Provision of preschool for children under 3 years old, availability of 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, in which a trained field worker interviewed the household members to obtain this information. All other variables are collected as part of the community survey, in which a trained interviewer collects the questionnaire from a community official, official records, and through direct observation.

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, the survey of supermarket and other modern food vendors were added in 2004. Ascertainment of cell phone ownership began in 2004, while that of 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, 11-13% for 1991-2000, and 4-5% 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 analyses of value-added by the scale, we utilize individual height and weight measurements. These measurements taken with portable equipment using standardized techniques by trained field workers. Body Mass Index (BMI) is calculated from weight in kilograms divided by height in meters squared. BMI is then used to categorize individuals as overweight or obese (BMI>25).

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

<u>Scale scoring</u>
We allot a maximum total of 10 points each to each of the 13 components. We had no *a priori* reason to justify weighting certain components more heavily than others, so we follow the scheme of Dahly and Adair who developed a similar scale to measure urbanicity in Cebu, Philippines (Dahly and Adair 2007).

Most components are constructed from multiple variables. 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). Appendix A provides details on the scoring algorithm used for each component.

Scale properties, reliability and validity

To evaluate whether the components operate at a unidimensional or a multidimensional scale, we first performed an exploratory factor analysis on the 13 components, without restriction to the number of factors estimated. We assessed dimensionality by number of factors with eigenvalues >1 and with a scree plot (Netemeyer RG 2003). After unidimensionality for one underlying construct was assumed, we assess 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 are reported(DeVellis 2003; Netemeyer RG 2003).

Temporal stability, also known as test-retest reliability, is 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 calculate these correlations for each consecutive survey wave, as well as between 1991 and 2006 surveys.

Criterion-related validity is typically assessed by comparing a scale's degree of agreement with a "gold standard" measurement of the construct the scale intends to measure (DeVellis 2003). In the case of measuring urbanicity, there is no accepted gold standard, so for criterion-related validity, we compare our scale to the existing official classification of urban status, using a dichotomous classification as rural or urban, as well as a categorical classification of urban city, county township, suburb, and village. This type of comparison is common place in other settings where their also is not an established gold standard; it is expected that correlations will be attenuated as compared to a comparison to a gold standard that was in fact a better reflection of the truth(Willett 1998).We assess criterion-related validity with Spearman's rank correlations using the 4-category measure. To compare our measure to the dichotomous urban/rural classifications, we dichotomize our measure based on receiver operator characteristic curve determination of maximized sensitivity and specificity for the year 2000, and compare the resulting categories using the Kappa statistic for agreement beyond chance.

Predictive validity, the extent to which the scale coincides with phenomenon known to differ by urban status(DeVellis 2003), will be measured by regression of urbanicity on per capita inflation-adjusted household income, as well as on odds of having more that one child. These are two phenomena that appear to be consistently related to urbanicity in China, and are therefore suitable regressors for an evaluation of predictive validity. We perform the regression using quintiles categories of the scale, modeled as indicators variables, as well the scale in its original continuous form.

To examine the value added of the scale, we look descriptively at the heterogeneity in the changes among the scale components across by urban status and by degree of total change in urbanicity as measured by the scale. We also compare estimates of incident overweight status by gradations of the new scale using logistic regression for a subset of people who had height and weight measured in both 1991 and 2004 and who were non-overweight in 1991.

Finally, we compare 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).

Results

The number of communities surveyed and the calender year for each wave are displayed in **Table 1**.

The thirteen components of the scale appear to represent a unidimensional underlying construct, which we assume to be urbanicity, as evidenced by the extremely high eigenvalue of only one factor in the exploratory factor analysis of the components. This is reflected in the scree plot for the factor analysis displayed in **Figure 1**.

Cronbach's α is very high for all years of the survey (α=0.85 – 0.90), indicating the scale has very good internal consistency (**Table 1**). Generally, an α of at least 0.70 is accepted as indicative of internal consistency (Netemeyer RG 2003). The corrected item-scale correlations, which calculate the correlation of each component with the remaining components in the scale, are also generally quite good; correlations >.40-.50 are desired (**Table 2**). Somewhat surprisingly, total population consistently has the lowest item-scale correlation among the pool of components each year, although, its correlation does increase over the survey years. Nevertheless, since the correlation of population with the rest of the scale does increase over time, and since it is an integral element for most definitions of urbanicity, it does not seem warranted to eliminate this component based on its initially low item-scale correlation, so it is retained.

Table 1. Cronbach's alpha for scale components by survey year

Year	Communities (N)	alpha
1989	190	0.85
1991	190	0.85
1993	188	0.86
1997	192	0.87
2000	217	0.86
2004	216	0.90
2006	218	0.90

Table 2. Individual component corrected item-scale correlations by survey year

	Year						
	1989	1991	1993	1997	2000	2004	2006
Component							
Population	0.25	0.23	0.26	0.28	0.41	0.52	0.52
Population Density	0.46	0.45	0.43	0.41	0.41	0.51	0.48
Education	0.60	0.62	0.61	0.63	0.67	0.66	0.68
Sanitation	0.69	0.71	0.70	0.72	0.74	0.75	0.77
Housing	0.67	0.73	0.71	0.74	0.82	0.80	0.80
Transportation	0.61	0.45	0.46	0.59	0.49	0.39	0.40
Communications	0.66	0.63	0.64	0.70	0.66	0.65	0.68
Health Infrastructure	0.60	0.68	0.67	0.63	0.50	0.61	0.62
Traditional Markets	0.52	0.43	0.43	0.46	0.43	0.55	0.62
Economy	0.66	0.74	0.73	0.64	0.68	0.68	0.69

Diversity	0.60	0.60	0.60	0.58	0.60	0.71	0.67
Modern Markets	0.43	0.45	0.45	0.50	0.59	0.73	0.75
Social Services	NA	NA	NA	NA	0.32	0.60	0.51

Further, the scale exhibits temporal stability in the test-retest reliability scenarios (**Table 3**). Correlations between each of the consecutive waves are quite high (r=0.90 to 0.94). Additionally, we'd expect the correlation between two more distant time points to be somewhat less than the correlations between two more proximal time points, and this is the case as the correlation between 1991 and 2006 scores is 0.84.

Table 3. Urban Scale Test-Retest Reliability between Survey Years

Comparison Years	Correlation
1989-1991	0.94*
1991-1993	0.91*
1993-1997	0.90*
1997-2000	0.91*
2000-2004	0.92*
2004-2006	0.93*

*p-value<0.001

There is evidence for the scale's criterion-related validity from the comparison to the official classification of communities as urban or rural. The Kappa statistics for agreement beyond chance of a dichotomized version of our scale and the "gold standard" range from fair to good inter-scale agreement (**Table 4**). Using our scale in its continuous form in comparison to the 4-category official classification of places, the high levels of Spearman rank correlations also support criterion-related validity (**Table 5**).

Table 4. Criterion-related validity: Kappa statistic for agreement beyond chance between dichotomized urban scale and official Chinese urban/rural classification

Year	Expected Agreement	Observed Agreement	Kappa Statistic
1989	63%	71%	0.21
1991	62%	71%	0.23
1993	61%	73%	0.31
1997	56%	69%	0.29
2000	52%	75%	0.47
2004	52%	77%	0.53
2006	51%	76%	0.50

Table 5. Criterion-related validity: correlation with official urban classification (village, suburb, town, city)

Year	Correlation
1989	0.78*
1991	0.77*

1993	0.78*
1997	0.75*
2000	0.78*
2004	0.77*
2006	0.77*

Spearman's rank correlation
P<0.001

The mean urban scores for communities increase over time as we would expect (**Table 6**). The linear and logistic regressions used to assess predictive validity indicate that our scale demonstrates good properties on this dimension, as well **(Tables 7** and **8**). Increasing scores on the urban scale are significantly associated with increases in adjusted per capita household income. Additionally, higher urban scores are associated with a significantly lower probability of having more than one child. Only women living in rural places are allowed by law to have more than one child, so the scale is associated in the direction that we would hope it to be.

Table 6. Mean urban scale scores by year and official Chinese classification

Year	Communities (N)	Village	Suburban	Town	Urban
1989	190	36(9.3)	52(13.8)	63(8.5)	66(9.0)
1991	190	38(10.4)	54(13.8)	64(8.9)	69(8.8)
1993	188	40(10.1)	59(13.4)	66(10.2)	71(7.6)
1997	192	44 (12.5)	60(14.9)	73(11.6)	77(9.8)
2000	217	48(11.1)	69(15.7)	81(9.7)	83(7.1)
2004	216	51(11.9)	74(18.8)	84(13.2)	92(7.7)
2006	218	53(12.6)	76(19.2)	84(13.4)	93(8.0)

Table 7. Cross-sectional linear regression of urban scale quintiles on adjusted household per capita

	Coefficient (Robust SE)		
Quintile	1991(n=3607)	2000(n=4313)	2006(n=4302)
Urban Scale Quintile 1(referent)	NA	NA	NA
Urban Scale Quintile 2	611 (181)**	826 (321)**	374 (626)**
Urban Scale Quintile 3	1045 (202)**	1600 (334)**	1117 (814)**
Urban Scale Quintile 4	1422 (226)**	2551 (385)**	5108 (665)**
Urban Scale Quintile 5	1813 (170)**	3781 (448)**	5499 (946)**

**p<0.01

Table 8. Cross-sectional logistic regression of urban scale quintiles on odds of having more than one child

	Odds Ratio (Robust SE)		
Quintile	1993 (n=2565)	2004 (n=2752)	2006 (n=2587)
Urban Scale Quintile 1(referent)	NA	NA	NA
Urban Scale Quintile 2	0.78 (0.22)	0.59 (0.11)**	1.16 (0.22)
Urban Scale Quintile	0.52 (0.13)**	0.59 (0.12)**	0.79 (0.17)
Urban Scale Quintile 4	0.36 (0.09)**	0.24 (0.06)**	0.26 (0.07)**
Urban Scale Quintile 5	0.11 (0.02)**	0.09 (0.02)**	0.13 (0.04)**

**p<0.01

Figures 2 and 3 display the heterogeneity of change in urbanicity component from 1991 to 2006. Overall change in score is calculated for communities that were included in the survey in both 1991 and 2006 (N=176). Of these communities 70% (N=123) are classified as rural according to the official classifications, while the remaining 30% (N=53) are classified as urban. Figure 2 represents the changes in urban score for communities that are classified as rural according to the official Chinese classifications. Figure 3 presents the same information for communities classified as urban by official classification. Communities are grouped by overall level of change in urban score, in categories of a decrease in urban score (urban score change <0), or an increase in score of 0 to 15, 15-30 points, >30 points.

Communities classified as urban had a higher overall mean change in score of this time period compared to those classified as urban (Urban mean change=25; Rural mean change=16; p-value<0.00001). Further, a larger percentage of the urban communities fell into the highest category of change (34% v. 12%), and a lower proportion of the urban communities fell into the lowest category of change (0% v. 6%). Of note, when evaluating growth or declines in the separate components we see that across community-type and across levels of urban score change, health infrastructure (health care facilities and pharmacies) appears to have consistently declined or experiences only very small growth. However, sanitation and housing infrastructure, which likely have large affects on individual health through basic hygiene pathways, consistently experienced fairly large gains.

Among the rural communities that experienced a decrease in urban score or a smaller level of increase, it appears that declines in scores are largely reflective of loss of both modern and traditional markets in these places, as well as losses in health infrastructure, and transportation infrastructure.

Figure 4 displays the predicted probabilities of incident overweight over the period 1991 to 2004 for the subset of the population who had BMI measured during both of these years (N=2,768), adjusted for gender (Figure 4 and Table 9). When comparing predicted probabilities only for the traditional dichotomous classification, the probabilities of the two groups are quite similar (0.27 v. 0.32; p=0.15) and the difference between them is not statistically significantly different. The variation that can be seen when by using the baseline and change in urban score from the new scale reveals one place where the new scale adds value over the traditional categorization.

Table 9. Logistic regression of incident overweight according to initial urban score and change in urban score (1991-2004)

	Coefficient	Robust SE	p-value
Baseline Urban Score in 1991	0.150	0.087	0.08
Urban Score Change	0.373	0.196	0.06
Baseline Urban Score in 1991 Squared	-0.001	0.001	0.17
Urban Score Change Squared	-0.008	0.004	0.03
Baseline Score*Score Change	-0.013	0.008	0.10
Baseline Squared*Score Change	0.0001	0.00006	0.14
Change Squared*Baseline Score	0.0003	0.0001	0.05
Baseline Squared*Change Squared	0.000003	0.000001	0.06
Gender	0.247	0.101	0.02
Intercept	-6.442	2.300	0.01

Discussion

Major social, demographic and economic change has occurred in both urban and rural areas across the developing world. China from 1980 to the present has recorded both some of the most rapid economic changes along with vast migration (Ma 2002). These changes have created dozens of cities with more than 5 million residents across the country, and also greatly affected the rural landscape. This paper has developed and tested the reliability and validity of a new scale to measure urbanicity on a continuum in China. Our scale uses >50 variables grouped into 13 components in an attempt to capture and distinguish the urban features of a place. We demonstrate unidimensionality of the scale, which is desirable since we aimed to tap the latent construct of urbanicity. The scale performs very well on tests of reliability and of content, criterion-related and predictive validity. Additionally, we briefly illustrate some applications in which the new scale adds valuable information to analyses of differences in adult health across levels of urbanicity.

To our knowledge, there are only a few published reports of scales that attempt to measure urbanicity on a continuum. McDade and Adair used a factor analysis approach to evaluate level of urbanicity for the Cebu Longitudinal Health and Nutrition Survey (McDade and Adair 2001). Subsequently, Dahly and Adair develop and test the reliability and validity of a new scale to measure urbanicity in Cebu, Philippines (Dahly and Adair 2007). The Cebu urbanicity scale was influential in the motivation for development of our scale described herein to measure ubanicity in China. The newly developed scale for China includes substantially more components of urbanicity compared to the Cebu urbanicity scale, and its reach spans a much larger geographic area and population. Additionally, the newly developed scale revisited the scoring algorithms of Dahly and Adair and revised them to provide finer gradations in component scores.

The analysis of incident overweight status is a brief example of the heterogeneity that may underlie many relationships between baseline and change in urbanicity of a place and health. The analysis is conducted among participants who were not overweight in 1991 and were followed up in 2004. The results examine expected probability of becoming overweight for individuals depending on their community's initial degree of urbanicity and its level of change in urbanicity over the follow-up period. We chose to model overweight as an outcome because it becomes a composite outcome of changes in food availability and intake and changes in activity level.

This analysis finds that for people living in communities with a low initial urban score (score ~30-40 in 1991), a higher level of change in urban score over the follow-up years is associated with a higher probability of become overweight during that time. However for people in communities with highest initial urban scores (scores~60-80), average probability of incident overweight is still higher than for less urban communities, but higher change in urban score does not predict higher probability of incident overweight. This relationship certainly deserves further exploration, but it could be posited that the type of urban change that occurs in already fairly urban places is less conducive to overweight status. Or, it could be that in highly urbanized environments, the obesity-promoting environmental changes have occurred and further increases in urbanicity do not affect behavior. This might be indicative of common changes in diet and activity or it might reflect quite different changes on diet (e.g. influenced by body image shifts so that urbanized youth want to adopt "Western thin body images") and activity. Further in-depth analysis is needed.

The scale developed in this paper has the potential for use in many analyses beyond this brief example. It can be used to further probe other health outcomes, consumption behaviors, demographic and economic activities, to name a few. After peer-review and publication, the scale and full documentation will be made publically available to researchers using the public use CHNS data.

The scale has many strengths, but limitations should also be noted. First, due to 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 change substantially. We also repeated these analyses excluding the component Traditional Markets, since, 1) out of any component, it had the highest frequency of imputed values, 2) and, conceptually, we were unsure whether this component was a good reflection of urbanization and urbanicity in more recent times. The reliability was essentially unchanged when traditional markets were 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.

Additionally, 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, and 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 (fast food, indoor restaurants and fixed food carts) could give a total possible of 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 includes 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 conceptualize the true levels of these as being close to nonexistent before we began collecting the data. Therefore we keep the total possible points consistent across years.

An additional limitation of the scale may be that total population has a fairly low corrected item-scale correlation. We use community population in our measure, rather than total city/town population. So for instance, the total population of one neighborhood in a city may not be consistently much larger than one neighborhood in a suburb or town. Accordingly, the population density measures display better correlations. We suspect that we would see better correlations between population and the rest of the scale had we used total city population, however, since the rest of the measures are at the community level, and since the results did not change when we down-weighted the population component (described below) we decided to retain this population measure.

Another potential limitation is the equal weighting of each of the 13 components. We decided to weight the components equally since we had no *a priori* reason to assign different weighting schemes to the different components. We did however test the influence of equal weighting by comparing the results to results from assigning different weights to each of the components in the total score. We performed a factor analysis of the 13 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 preformed the equivalent criterion-related validity testing for the scores from factor analysis derived weights. We found virtually the same results with the two methods.

In summary, the urbanicity scale developed for China captures enormous changes in context across time and space. It performs very well on all conducted tests of internal consistency, reliability, content validity, criterion-related validity and predictive validity. The overall scale and many of its components can be used to examine an array of outcomes. As shown in this analysis, it clearly predicts incident obesity and weight dynamics in China and promises to be most useful for other economic, demographic, social welfare and health outcomes.

Acknowledgment

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. We wish to thank Jim Terry for the tremendous programming support, without which this project would not have been possible.

References

Batal, M., C. Boulghourjian, et al. (2006). "Breast-feeding and feeding practices of infants in a developing country: a national survey in Lebanon." Public Health Nutr **9**(3): 313-9.

Brockerhoff, M. (1995). "Child survival in big cities: the disadvantages of migrants." Social Science & Medicine **40**: 1371-83.

Brockerhoff, M., Brennan, E. (1998). "The Poverty of Cities in Developing Regions." Population and Development Review **24**: 75–114. .

Dahly, D. L. and L. S. Adair (2007). "Quantifying the urban environment: a scale measure of urbanicity outperforms the urban-rural dichotomy." Soc Sci Med **64**(7): 1407-19.

DeVellis, R. (2003). Scale Development: Theory and Applications. Thousand Oaks, CA, Sage Publications.

Fotso, J. C. (2006). "Child health inequities in developing countries: differences across urban and rural areas." Int J Equity Health **5**: 9.

Hinde, P. R. and A. J. Mturi (1996). "Social and economic factors related to breast-feeding durations in Tanzania." J Biosoc Sci **28**(3): 347-54.

Hu, X., S. Cook, et al. (2008). "Internal migration and health in China." Lancet **372**(9651): 1717-9.

Ma, L. J. C. (2002). "Urban transformation in China, 1949-2000: a review and research agenda." Environment and Planning A **34**: 1545-69.

McDade, T. W. and L. S. Adair (2001). "Defining the "urban" in urbanization and health: a factor analysis approach." Social Science & Medicine. **53**(1): 55-70.

Mendez, M. A., C. A. Monteiro, et al. (2005). "Overweight exceeds underweight among women in most developing countries." Am J Clin Nutr **81**(3): 714-21.

Mendez, M. A., C. A. Monteiro, et al. (2005). "Overweight exceeds underweight among women in most developing countries." American Journal of Clinical Nutrition **81**(3): 714-721.

Mendez, M. A. and B. M. Popkin (2005). "Globalization, urbanization, and nutritional change in the developing world." Electronic Journal of Agricultural and Development Economics **1**: 220-241.

National Research Council (2003). Cities Transformed: Demographic Change and Its Implications in the Developing World. M. M. Panel on Urban Population Dynamics, Stern R, Cohen B, Reed HE. Washington, DC, Committee on Population, Division of Behavioral and Social Sciences and Education.

Netemeyer RG, B. W., Sharma S (2003). Scaling Procedures: Issues and Applications, Sage Publications.

Popkin, B. M., R. E. Bilsborrow, et al. (1982). "Breast-feeding patterns in low-income countries." Science **218**(4577): 1088-93.

Poston, D. L. and M. X. Mao (1998). "Interprovincial migration in China, 1985-1990." Res Rural Sociol Dev **7**: 227-50.

Rondinelli, D. A. (1983). Secondary Cities in Developing Countries: Policies for Diffusing Urbanization. Beverly Hills, CA: , Sage Publications.

Rondinelli, D. A. (1987). "Cities as agricultural markets. ." The Geographical Review **77**(4): 408-20.

United Nations (2002). World Urbanization Prospects, 2001. United Nations.

Van de Poel, E., O. O'Donnell, et al. (2007). "Are urban children really healthier? Evidence from 47 developing countries." Soc Sci Med **65**(10): 1986-2003.

Vlahov, D. and S. Galea (2002). "Urbanization, urbanicity, and health." J Urban Health **79**(4 Suppl 1): S1-S12.

Willett, W. (1998). Nutrition Epidemiology Second Edition. New York, Oxford University Press.

Yang, X. and F. Guo (1999). "Gender differences in determinants of temporary labor migration in China: a multilevel analysis." Int Migr Rev **33**(4): 929-53.

[i] For further information and articles that have used the CHNS data, see http://www.cpc.unc.edu/projects/china.

Figure 1. Scree plot of eigenvalues after exploratory factor analysis

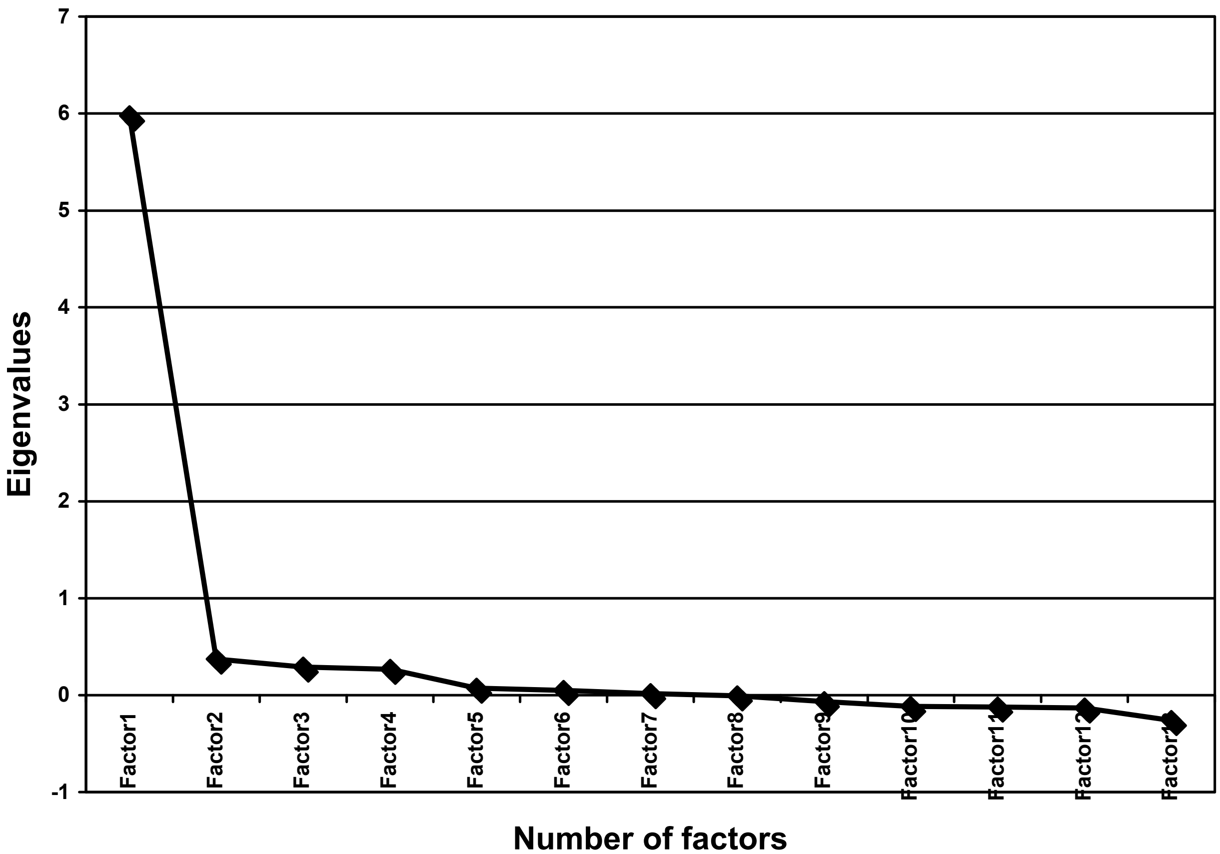


Figure 2. Change in individual components by level of overall change in urban score for “Rural” Communities (1991-2006)

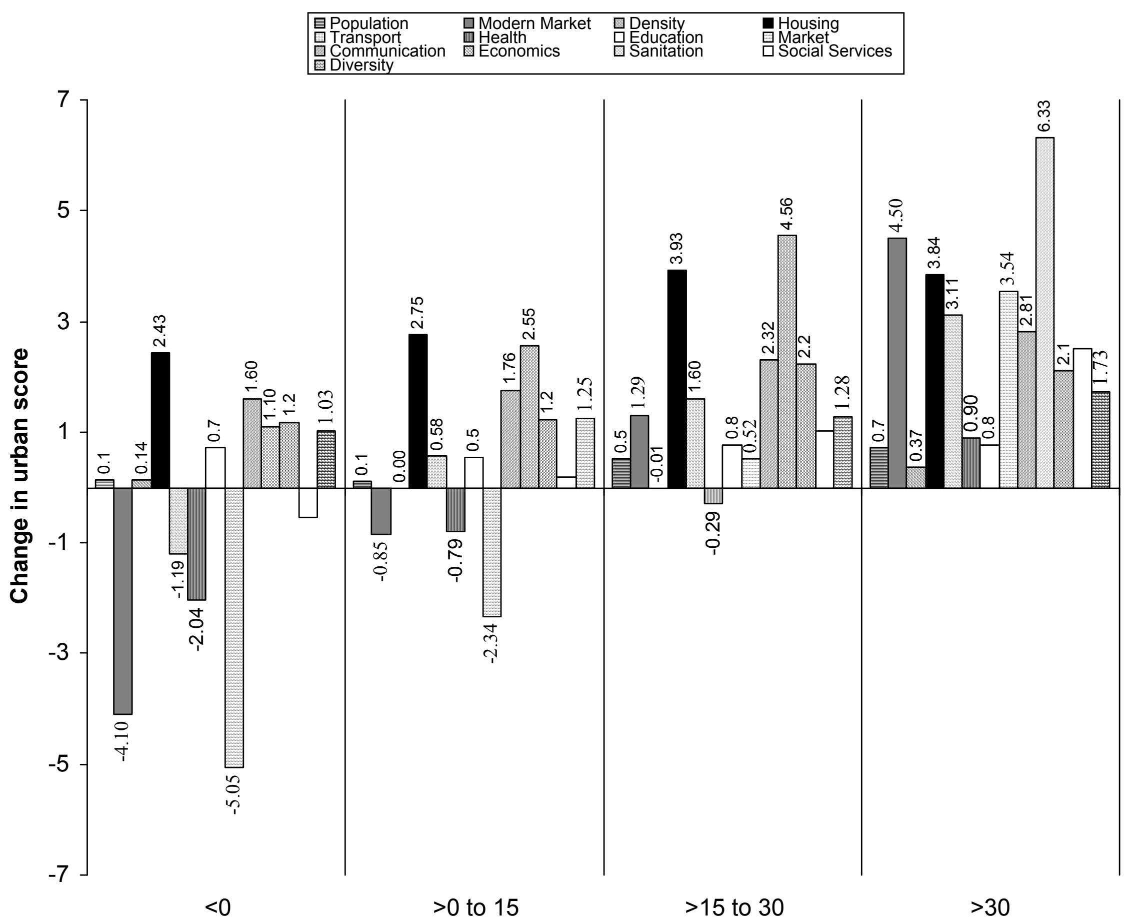


Figure 3. Change in individual components by level of overall change in urban score for "Urban" Communities (1991-2006)

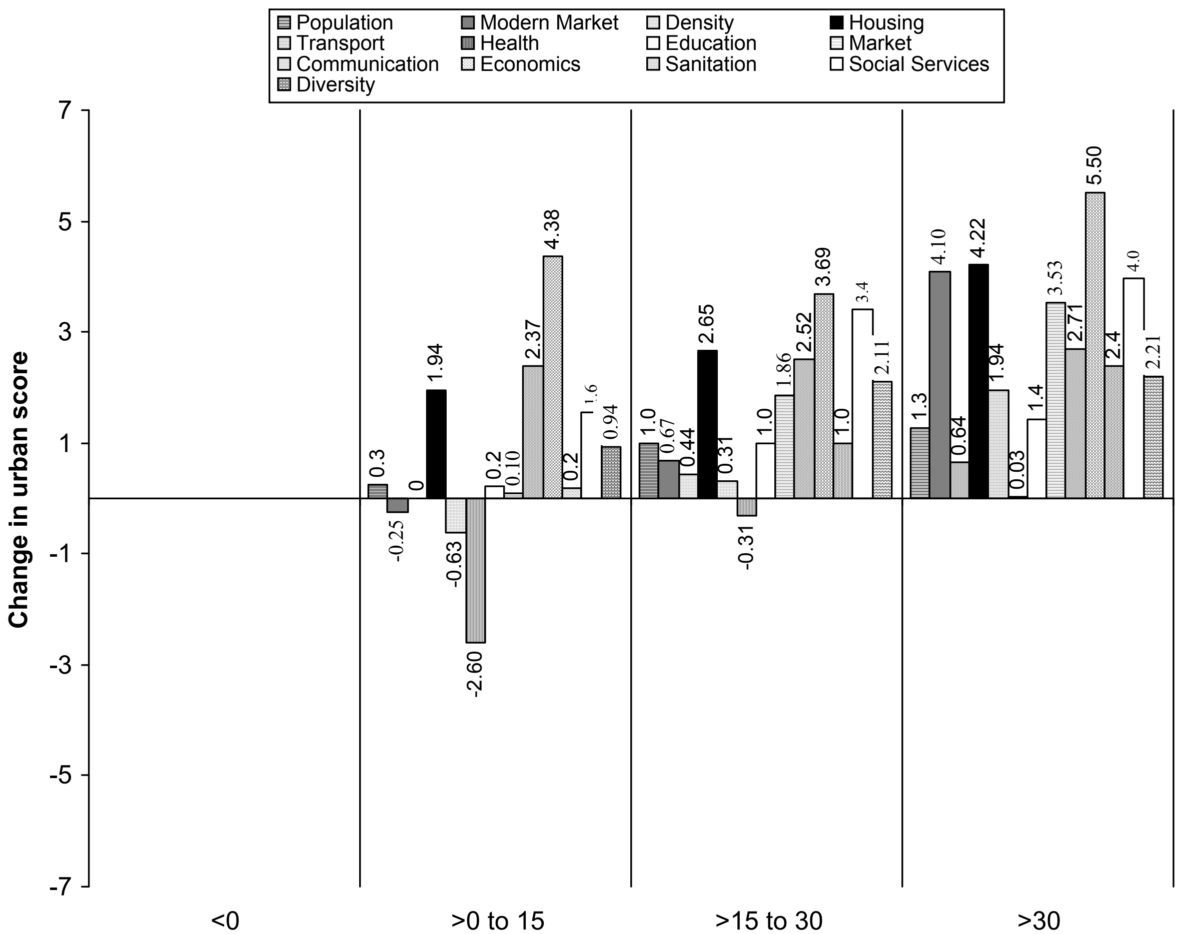
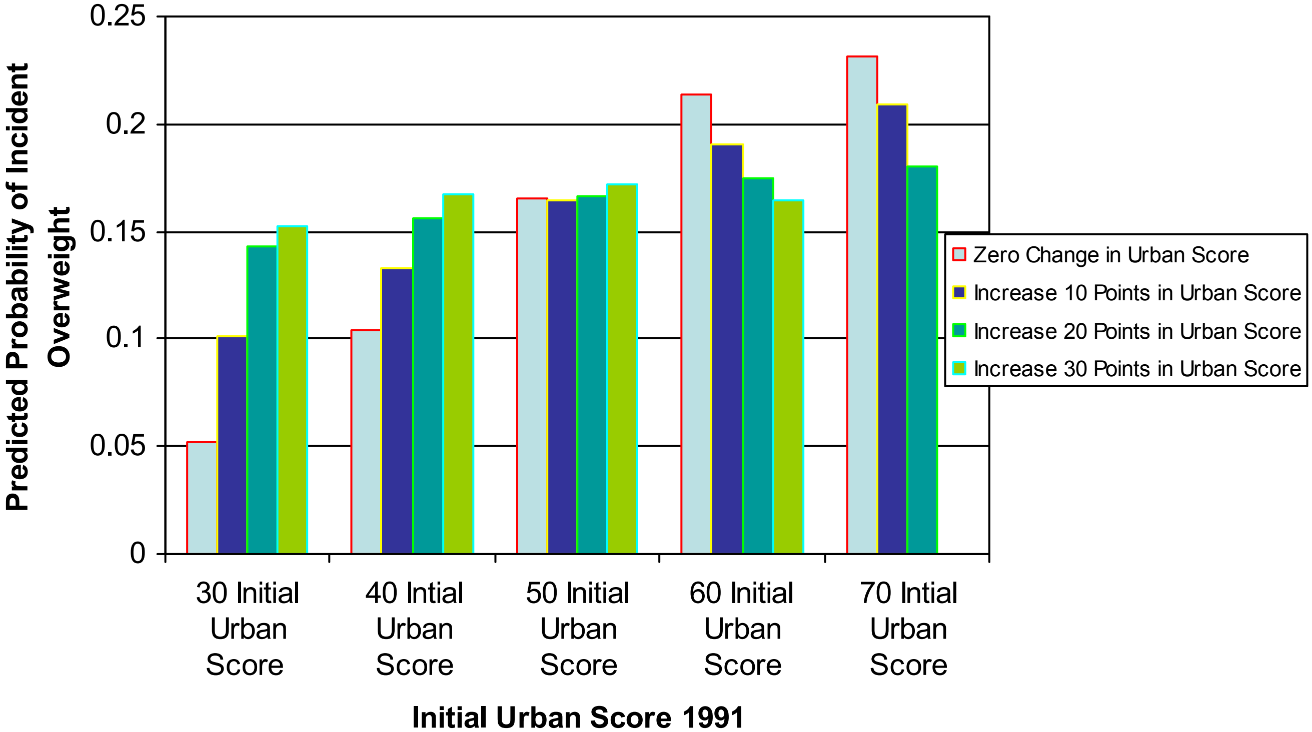


Figure 4. Predicted probability of incident overweight (1991-2004) by change in urban score over follow-up for varying levels of initial urban score.



Appendix A

Component	Scoring scheme
Community Population Size	(see table below)

Points	Log Population Range		Population Range	
0	4.8	5.13	0	169
0.5	5.13	5.46	169	235
1	5.46	5.79	235	327
1.5	5.79	6.12	327	454
2	6.12	6.45	454	632
2.5	6.45	6.78	632	880
3	6.78	7.11	880	1224
3.5	7.11	7.44	1224	1702
4	7.44	7.77	1702	2368
4.5	7.77	8.1	2368	3294
5	8.1	8.43	3294	4582
5.5	8.43	8.76	4582	6374
6	8.76	9.09	6374	8866
6.5	9.09	9.42	8866	12332
7	9.42	9.75	12332	17154
7.5	9.75	10.08	17154	23860
8	10.08	10.41	23860	33189
8.5	10.41	10.74	33189	46166
9	10.74	11.07	46166	64215
9.5	11.07	11.4	64215	89321
10	11.4	11.73	89321	124243

Population Density (people per km2)

Points	Log Population Range		Population Range	
0	0	0.58	1	1
0.5	0.58	1.16	1	3
1	1.16	1.74	3	5
1.5	1.74	2.32	5	10
2	2.32	2.9	10	18
2.5	2.9	3.48	18	32
3	3.48	4.06	32	57
3.5	4.06	4.64	57	103
4	4.64	5.22	103	184
4.5	5.22	5.8	184	330
5	5.8	6.38	330	589
5.5	6.38	6.96	589	1053
6	6.96	7.54	1053	1881
6.5	7.54	8.12	1881	3361
7	8.12	8.7	3361	6002
7.5	8.7	9.28	6002	10721
8	9.28	9.86	10721	19148
8.5	9.86	10.44	19148	34200
9	10.44	11.02	34200	61083
9.5	11.02	11.6	61083	109097
10	11.6	12.18	109097	194852

Economic
Variables used:
- Real wage for ordinary male worker
- Percent of community engaged in nonagriculture

If male wage>16.9 (real, adjusted by CPI), then econ score = 2 * (proportion not in agriculture))/20

If male wage<16.9(real, adjusted by CPI) then econ score = 1*(proportion not in agriculture))/20

Traditional Markets
Variables used:
- Types of goods/markets available (grains, oil, meat, vegetables, fish, bean curd, milk, fabric and fuel);
- Distance to markets; number of days markets are open;

First, 1.1 points given for each of nine stores in or near community. Then, points=((10* #stores in*#days open)/70) + ((5*#stores near*#days open near)/70). This would give a score of 10 if you had all stores open on all days of the week in your community and a score of 5 if all stores were open 7 days in a nearby

Supermarket and food vendors

Variables Used:

- Supermarket/Hypermarket in neighborhood
- How many of each of the following establishments in neighborhood: indoor food restaurant, fixed outdoor food stall, fast food restaurant, ice cream parlor, bars, cafes, internet cafes, fruit and vegetable stores, bakeries.

Fixed Outdoor food stalls (variable o61) asked starting in 1989.
Mobile Food Carts (variable o62) asked starting in 1989
Indoor Restaurants (variable o60) asked starting in 1989
Fast food restaurants (variable o276) asked starting in 2004
Ice Cream Parlors (variable o279) asked starting in 2004
Fruit and vegetable vendors (variable o280) asked starting in 2004
Bakeries (variable o277) asked starting in 2004
Cafes (variable o281) asked starting in 2004
Internet cafes (variable o282) asked starting in 2004

Supermarkets use variable o285, which started in 2004 and asks how many supermarkets are within 30 min bus ride. Categorize this into 3 categories: 0 supermarkets, 1 supermarket and >=2 supermarkets. Then give for "supermarket" give points as follows:

Supermarkets	Points
0	0
1	1
ge 2	2

For "vendors_old": For each community in each year, add together the values of o60, o61, o62 to get a total score for the vendor info collected 1989-2006. To give a score for the "vendors_old" portion, assign points according to the following scale:

Number of Vendors	Points
0	0
1 to 2	1
gt 2 to 4	2
gt 4 to 7	3
gt 7 to 10	4
gt 10 to 20	5
gt 20 to 50	6
gt 50 to 100	7
gt 100	8

For "vendors_new": For each community in each year, add together the values of o276, o277, o279, o280, o281, o282 to get a total score for vendor info collected 2004-2006. To give a score for the "vendors_new" portion, assign points according to the same scale as old vendors

For 2004-2006, for total "modern market" score: add points from supermarket + ((points from vendors_old +points from vendors_new)/2))
For 1989-2000, for total "modern market" score: add points from supermarket + (points from vendors_old)

Education
Variables used:
Average highest education attainment among adults.

Take actual community average and multiply it by 10/6 to scale the whole range to 10.
0-6: 0=none, 1=graduated primary school, 2=lower middle school, 3=upper middle school, 4=middle, technical or vocational, 5=3 or 4 year college degree, 6=masters degree or higher.

Diversity
Variables used:
- Community variance in education level
- Community variance in income level

Education: Created a variable that is the variance in education for each community each year.

Education Variance (comm_var)	Points
0-0.50	0
>0.50-0.75	1
>0.75-1.0	2
>1.0-1.25	3
>1.25-1.5	4
>1.5-1.75	4.5
>1.75-2.0	5
>2.0-2.25	6
>2.25-2.50	7
>2.50-2.75	8
>2.75-3.00	9
>3.00	10

Income Diversity: Created a variable that is the variance in income for each community in each year, divided by 1000 for more manageable numbers. (Scale based on log scale for huge range in numbers).

Income Variation (comm_var1000)		Points
From	To	
>0	20	1
>20	90	2
>90	400	3
>400	1800	4
>1800	8100	5
>8100	36300	6
>36300	162750	7
>162750	729400	8
>729400	3269000	9
>3269000		10

For total score, take mean of Education Diversity and Income Diversity scores.

Health Variables used: • Type of health clinic • Distance to health clinic • Number of health clinics in neighborhood	**Give points according to scheme for highest ranking health facility:** **If highest ranking health facility is:** Village clinic or Private Clinic or "other clinic if in R" (1, 14, 19 if in rural) = give 1 point Township hospital or TWP family planning service or "other hospital if in R" (2, 21, 20 if in rural)= give 2 points County hospital or county MCH or MCH clinic or work unit clinic (3, 22, 13, 6)= give 3 points City hospital or neighborhood clinic or community hospital, or district hospital or army hospital or Specialty hospital or private hospital or "other clinic if in urban" or other hospital if in urban" or city MCH or workers hospital (9, 4, 5, 8, 10, 11, 15, 19 if in urban, 20 if in urban, 23, 7)=give 4 points **Then give** An additional 1 point for pharmacy/drug store in community only (i.e. not nearby) **Then** Let total possible from above equal 8 points only **Then** Adjust up if they have more than one clinic: multiply their score by 1.25 so that their total can equal 10 **For communities with only a health facility in a neighboring community or >12km away**, Adjust score down by multiplying score by 0.5. .
Transportation Variables Used: • Road type (paved, gravel, dirt) • Availability of bus stop or train station • Distance to bus or train station	Type of road: dirt=0, stone/gravel/mixed=1, paved=2 Bus station: in community=2, close, but not in community =1, no close bus = 0. Train station: in community=2, close, but not in community =1, no close bus = 0. Scaled to 10 points = (road + bus + train)*1.6666

Housing Variables Used: • Prevalence of electricity • Prevalence of indoor plumbing • Prevalence of natural gas	Electricity timing points = hours per day of electricity/24 * days on/7 days)*10 In house tap water= percent of community with indoor piped water/10 In house flush toilet = percent of community with flush toilet/10 Cooking with gas=percent of community that cooks with gas/10 For score take the mean of electricity, water, toilet and gas.
Sanitation Variables Used: • Water treatment plant available to household; • Absence of excreta in public space	Prevalence of households with treated water/10 Prevalence of households without excreta present outside/10 For score take mean of above 2 items.
Communications Variables Used: • Percent of households with a television; • Percent of households with a computer; • Percent of households with a cell phone; • Cinema in neighborhood; • Newspaper available to the neighborhood; • Telephone service available to community; • Postal service available to community;	Scoring: Percent of households with color TV/10 Percent of households with computer/10 Percent of households with cellphone/10 Take mean of above 3 numbers and divide by 1.66666 (or 10/6). Then Add 1 point for each of the following: Cinema in neighborhood Newspaper available Postal service available Telephone service available
Social Services Variables Used: • Child care center (0-3yrs old) • Type of insurance offered in community	2.5 points given for each of the following: A child care center for <3year olds; commercial insurance; free medical insurance; women and children insurance. If childcare center is not in village/neighborhood, but is in a neighboring village less than 25 km away, give only 1.25 points for this.