Research Article

Estimating rural–urban disparities in self-rated health in China: Impact of choice of urban definition

Audrey Dorélien

Hongwei Xu

Contents

Estimating rural–urban disparities in self-rated health in China: Impact of choice of urban definition

Audrey Dorélien[1]

Hongwei Xu[2]

Abstract

OBJECTIVE AND METHODS
We use the 2014 China Family Panel Studies (CFPS), which includes anchoring vignettes, to provide an up-to-date assessment of rural–urban health disparies as measured by self-rated health (SRH) in China. Our analysis is based on multiple definitions (*hukou* and the two different residence-based definitions) of rural–urban and migration status; previous research was inconclusive due to the use of different definitions and concerns about status-based differential health expectations (reporting heterogeneity).

RESULTS AND CONCLUSIONS
We find a nonlinear difference between rural and urban Chinese in how they self-assess health status, regardless of the urban definition used. Urban respondents do not always hold a higher standard for self-assessment of health. Instead, their rating styles depend on the level of latent health. After controlling for the reporting heterogeneity, we find on average a slight urban advantage in SRH, but it is most pronounced when using the statistical (density dependent) definition of urban.

CONTRIBUTION
We study rural–urban health disparies based on three different urban definitions and migration status. Although we examine the urban definitions that are specific to China, we demonstrate a mindful approach when multiple definitions exist and caution against any simplistic approach that ignores context-specific urban definition. We also provide clear illustrations of the different types of reporting heterogeneity, as well as a way to visualize the cut-points, thresholds, and latent health estimates.

---

[1] University of Minnesota Twin Cities, USA. Email: dorelien@umn.edu.
[2] City University of New York, USA. Email: hongwei.xu@qc.cuny.edu.

1. Introduction

Health disparities between rural and urban populations have become an increasingly important public health concern in developing countries where the world's largest wave of urbanization is occurring (UN 2015). Historically, urban areas were associated with a health penalty as overcrowding, pollution, and concentrated poverty contributed positively to the prevalence of infectious diseases in cities. As city living standards improved, the urban health penalty faded and generally gave rise to an urban health advantage in medical care utilization and life expectancy (Woods 2003). China was no exception; the disease environment was worse in urban than in rural areas until the early twentieth century when a variety of public health and sanitary measures were first introduced in cities (Campbell 1997). The urban health advantage that first appeared in the first half of the twentieth century slightly diminished as preventive health care became widely available in rural areas from 1949 to the late 1970s (Zhang and Kanbur 2005). The trend of narrowing the rural–urban health gap was halted, if not reversed, after the nationwide collapse of the rural cooperative medical program in the 1980s caused by the market-oriented reform (Xu and Short 2011).

Being home to 20% of the world's urban population (UN 2015), China's rapid urbanization – and its associated economic growth and rising inequality – represents a large-scale process that has occurred around the world. How will the rapid urbanization shape the pattern of rural–urban health gap in China? On one hand, urbanization is generally associated with economic growth, better education, more job opportunities, and higher income – all socioeconomic determinants of health. Urbanization is also likely to improve access to quality health-care services, which contributes positively to the urban health advantage. On the other hand, rapid urbanization in developing countries can have negative consequences for population health due to increased pollution, overcrowding, and the adoption of lifestyles characterized by an unhealthy diet and reduced physical activity (Van de Poel, O'Donnell, and Van Doorslaer 2009, Van de Poel, Van Doorslaer, and O'Donnell 2012). These contradicting forces stimulate a growing research interest in understanding China's rural–urban gap in health.

Capitalizing on newly available nationally representative data, this study provides an up-to-date assessment of health disparity (as measured by self-rated health) by rural–urban and migration status, perhaps the most fundamental marker of social stratification and inequality in China. A key innovation is that we do this by analyzing multiple definitions of urban and migration status and assess to what extent reporting heterogeneity may influence the results. We find a slight urban advantage regardless of urban definition used. We also find nonlinear differences in how rural and urban Chinese rate their health; rural Chinese are more likely to choose extremes (e.g., rate their health as poor or excellent compared to urban Chinese). Although we examine the urban

definitions that are specific to China, we demonstrate a mindful approach when multiple definitions exist and caution against any simplistic approach that ignores context-specific urban definition. We focus on self-rated health, but our analytical strategy can be easily applied to other health indicators.

2. Background

2.1 Urban definitions in China

The empirical findings on rural–urban differences in health in China are far from conclusive, which we suspect is partly attributable to the lack of transparency and consistency in the choice of urban definition. Defining urban and rural areas is not a trivial task. Urbanness is multidimensional with respect to human settlements, encompassing ecological, political, social, economic, and cultural components (Dorélien, Balk, and Todd 2013). Government agencies and social scientists often simplify the rural–urban classification based on readily available statistics such as population size, administrative boundaries, and proportion of adult population in nonagricultural occupations (Utzinger and Keiser 2006). But even when urban areas are defined by uniform criteria such as population size and densities, the thresholds incorporated into datasets vary substantially from one country to another. Secondary data analyses of rural–urban health disparities often take for granted the rural–urban definitions provided by the data collection team, ignoring alternative definitions and their empirical implications.

Defining city population and urban status is particularly difficult in China, where three official definitions exist (Chan 2007). An area can be defined urban based on administrative definition: A community is classified as urban if it is governed by an urban residential committee and rural if a rural village committee governs it. Administratively defined urban cities (*jianzhi shi*) receive more financial support from the state, which can be used for urban services (Ma and Cui 1987; Qin and Zhang 2014). Consequently, individuals living in communities with a nonurban administrative designation might have less access to government services compared to individuals living in a community of similar population size and density but with an administrative urban designation. According to the statistical definition from the National Bureau of Statistics of China (NBS), a community is treated as urban if its population density exceeds 1,500 persons per square kilometer and rural otherwise. Urban statistical areas can be found within urban administrative areas and rural administrative areas (Chan 2007). In China, urban status is also conferred at the individual level through hukou designation: An individual is considered as urban if he/she possesses a nonagricultural hukou and rural if he/she possesses an agricultural hukou. Furthermore, hukou status is specific to an

administrative unit and entitles holders to receive welfare and services while residing in that unit. Consequently, residents of the same neighborhood who have different hukou statuses might have unequal access to services. In this study, we systematically examine China's rural–urban health gap across the three urban definitions: one definition at the individual level (hukou status) and two definitions at the community level (statistical and administrative classifications of urban versus rural communities).

2.2 Migrants

We also improve upon the literature by incorporating rural-to-urban migration into the research on rural–urban health gap. Similar to many other developing countries, a driving force behind China's urbanization is the large inflow of rural-origin population into existing urban settlements, looking for nonagricultural job opportunities (Shen 2006). Rural-to-urban migration accounts for nearly half of China's urban population growth between 1980 and 2010 (Chan 2013). In the rural–urban health literature, migrants to cities tend to be categorized as part of China's urban population to maintain a simple dichotomous comparison. In the migration literature, a growing number of studies have assessed the health consequence of migration (Chen, Chen, and Landry 2013; Qiu et al. 2011; Song and Sun 2016) and the possibility of selective migration by health (Chen 2011; Tong and Piotrowski 2012). However, these studies usually define migrants by comparing an individual-level hukou status with a given, often unspecified, community-level urban status. In this study, we define migrants in two different ways given two community-level urban definitions available. Therefore, we evaluate whether different community-level urban definitions and thereby different definitions of migration status affect the estimate of a migrant's health status.

2.3 Self-rated health

We chose self-rated health (SRH) status as the main health indicator. For its wide availability in household survey data and robust predictive power for mortality and morbidity (Benjamins et al. 2004; Idler and Benyamini 1997; Singh-Manoux et al. 2006), SRH remains a popular health measure to date not only in developing countries but also in developed countries (e.g., Do, Frank, and Iceland 2017; Zajacova, Huzurbazar, and Todd 2017).

2.4 Previous research on rural–urban differences in SRH in China

From the existing literature, it is not clear whether there is an urban advantage or disadvantage in SRH in China. Using data from the Survey on China's Support Systems for the Elderly in Rural and Urban China, Zimmer and Kwong (2004) find no differences in SRH when comparing urban and rural elders. They used the administrative urban and rural definitions in their study, and their SRH response categories were dichotomized into 1 = healthy, and 0 = fair or unhealthy. The authors also mentioned that one of the factors behind their findings might have been that urban elders may be more disposed to report health conditions than are rural elders.

An urban disadvantage in SRH has been documented in China's general population. Using data from the 2004 National China Survey on Attitudes towards Inequality and Distributive Justice and treating SRH on a 5-point Likert scale as an interval measure, Whyte and Sun (2010) find that rural villagers reported better health status than did urban citizens after controlling for other socioeconomic and demographic characteristics. However, by looking at their raw data without any controls they obtained different results – for instance, rural residents are more likely to rate their health as good or better but are less likely to rate their health as okay or better compared to urban citizens.

Based on data from the Chinese Health and Nutrition Survey (CHNS), Pei and Rodriguez (2006) report that rural respondents were less likely than urban respondents to report poor or fair health versus good or excellent health after controlling for individual characteristics and provincial income inequality. They used the administrative urban and rural definitions in their study. Also, using CHNS data, van De Poel and colleagues (2012) go a step further and find that urbanization (an increase in the urbanicity index they created) was associated with a decline in self-rated health in the short run. The authors raised the point that health status might not have worsened with increased urbanization, but rather residents in increasingly urbanized locations might have raised their health expectations. As with SRH among the elderly, this highlights one of the limitations of self-rated health measures – heterogeneities in health ratings.

It is worth noting that much of the existing literature doesn't simply report the associations between rural–urban status and SRH but also the association after controlling for many factors. Some of these control variables are necessary to adjust for rural–urban difference in basic demographic compositions (e.g., age and sex). Other covariates (e.g., education, income, and health-care provision), however, are potential mediators of the effect of urban residence on health. Controlling for these covariates will obscure the unconditional association between rural–urban status and health.

2.5 Reporting heterogeneity

Several studies have suspected that reporting heterogeneity may explain a surprisingly negative association between socioeconomic status and SRH in China and other developing countries such as Thailand and the Philippines (Luo and Wen 2002; Pei and Rodriguez 2006; Whyte and Sun 2010; Zimmer and Amornsirisomboon 2001; Zimmer et al. 2000). Following recent studies using the same data (Xu and Xie 2016, 2017), we employ the anchoring vignette method to examine to what extent reporting heterogeneity may (or may not) explain the observed rural–urban gap in SRH.

Most of the previous studies of SRH in China do not control for reporting heterogeneity, but all state that it is a likely concern (Bago d'Uva et al. 2008; van De Poel, Van Doorslaer, and O'Donnell 2012; Whyte and Sun 2010; Zimmer and Kwong 2004). Instead, they assume, albeit implicitly, that individuals from different social backgrounds employ the same standard in their subjective health ratings. In reality, however, individuals of different social groups are likely to have had different health experiences and may thus vary in their subjective ratings of the same objective health status. For example, respondents of high socioeconomic status (SES) may compare themselves to their better-off peers and hence adopt a higher standard for what is considered "excellent" health than do those of low SES or vice versa. This could then result in an inflated level of self-rated health among low-SES respondents compared to high-SES respondents, despite the high-SES group's advantage in true health status (Dowd and Todd 2011; Schnittker 2005). This measurement issue is also known as reporting heterogeneity, or differential item functioning (King et al. 2004).

Reporting heterogeneity can follow two patterns, parallel cut-point shift and nonparallel cut-point shift (Figure 1). In the former case, the same covariate affects all the cut-points in the same way. As illustrated in the top panel of Figure 1, for example, we expect urban respondents to adopt higher thresholds in evaluating their health in that their cut-points shift up in parallel for each health category compared with rural respondents. In the latter case, the association between a covariate and the corresponding cut-point shift is nonlinear. For example, as shown in the bottom panel of Figure 1, urban respondent may employ higher cut-points for the ratings of very good (versus good) or excellent (versus very good) health but a lower cut-point for the rating of good (versus fair) health.

Figure 1: **Visual representation of the different types of group reporting heterogeneity. Breaks in the colors represent the location of the cutpoints**
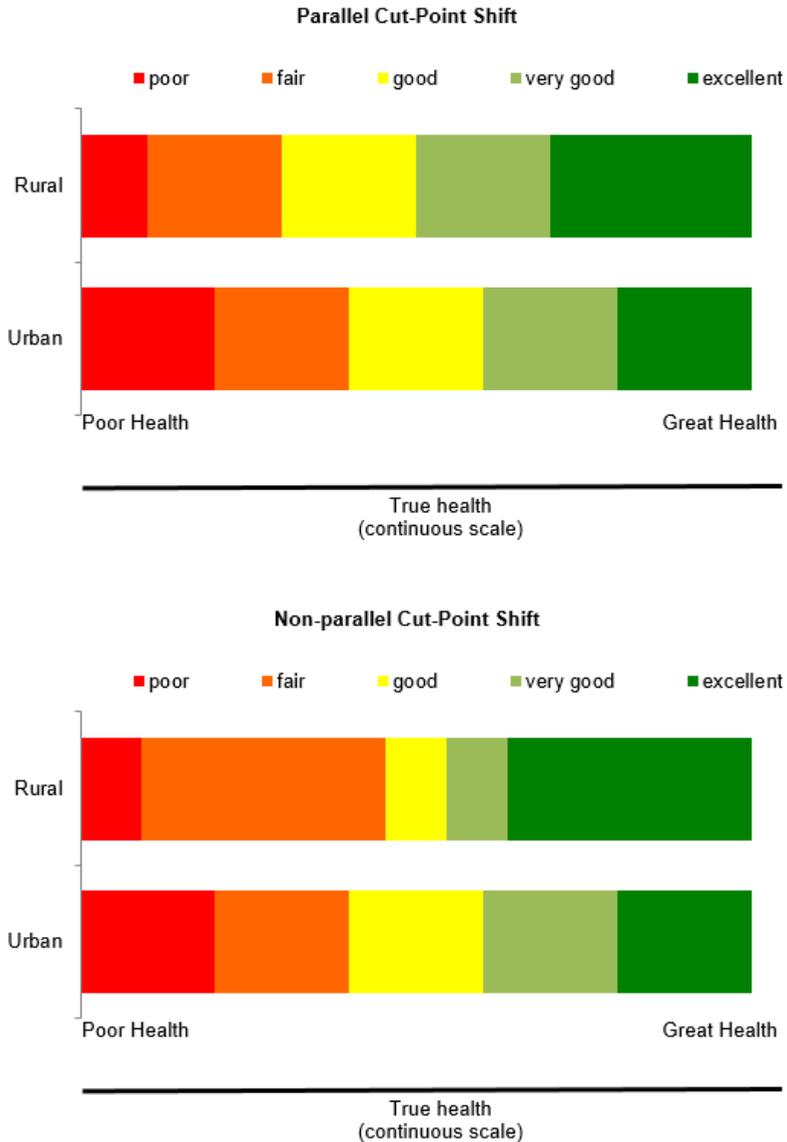
Anchoring vignettes help address the problem of heterogeneities in health ratings – systematic differences in health reporting (Grol-Prokopczyk, Freese, and Hauser 2011). Anchoring vignettes are brief texts describing hypothetical individuals with varying levels of health; survey respondents are asked to rate the health status of the hypothetical individual. The survey respondents are asked to assume that the hypothetical individuals described are of the same age and background as themselves, and to evaluate that person's health as if they were rating themselves (response consistency). We assume that the vignettes are perceived in the same way and on the same unidimensional scale by all respondents (vignette equivalence assumption; King et al. 2004). If those two assumptions hold, we can apply the estimated individual-specific cut-points from the vignettes data to anchor respondents' ratings of their own health status. More technical details about the vignette methodology can be found in King et al.'s seminal paper (2004).

There is evidence of reporting heterogeneity in China. Drawing on data from the 2000–2001 WHO Multi-Country Survey Study on Health and Responsiveness, Bago d'Uva et al. (2008) report that in general, Chinese respondents of different demographic (age, sex, and rural/urban residence) and socioeconomic status (education and income) adopted systematically different reporting styles in rating their health status, although the statistical significance of reporting heterogeneity varied by health domains (i.e., mobility, cognition, pain, self-care, usual activities, and affect). Drawing on data from the 2012 wave of the China Family Panel Studies, Xu and Xie (2016, 2017) find that a variety of demographic (age, sex, marital status, rural/urban residence, and migration) and socioeconomic (cognition, education, income, and political capital) characteristics were associated with cut-point shifts in self-rated general health and that cut-point shifts were mostly nonparallel. Xu and Xie (2016) also demonstrate that the anchoring vignettes implemented in the CFPS could effectively correct for reporting bias in SRH.

2.6 Research hypotheses

The research background reviewed above leads us to make several predictions. Compared with their rural counterparts, we hypothesize that urban respondents are more aware of new disease risks and hold a higher health expectation, and thus underreport their health status, resulting in an underestimated urban advantage in SRH. We expect to obtain a larger estimate of urban health advantage after adjusting for reporting heterogeneity.

We further expect such a gap to be more evident when we choose the statistical rather than the administrative urban definition. The reason is that the statistical definition captures the urbanization process in a timely manner whereas the administrative definition tends to lag behind. As a rural community experiences industrialization and growth in its population size and density, it will be quickly reclassified by China's

National Bureau of Statistics as an urban community for statistical purpose. In contrast, it will take a much longer time to convert the government administrative structure from a rural village committee to an urban residential committee, a transition that involves multiple institutional changes, including dismantling the land-use right collectively possessed by the villagers, converting the hukou status of every village from agricultural to nonagricultural, and transferring various social welfares (e.g., health insurance program and pension) from rural to urban. In short, it is more cumbersome for a rural community to be redesignated as an administratively defined urban community despite its urbanized socioeconomic and health environments.

On the other hand, hukou is an indicator of individuals' social status and functions to deny the rural population the access to urban-biased public goods and social welfare benefits. Even when they migrate to work and live in urban areas as migrant workers, rural hukou holders are treated as second-class citizens (Pan, Lin, and Peng 2012). Nevertheless, rural-to-urban migrants are found to be positively selected on the basis of SRH, although the health-selection effect has diminished during the recent decade (Tong and Piotrowski 2012). Taken together, if considering residence-based urban definition (i.e., statistical and administrative definitions) only, we should expect a rural disadvantage in SRH, again after adjusting for reporting heterogeneity. Part of the rural disadvantage may be attributed to the difference in the demographic composition. That is, the urban residents consist of higher-status urban natives and the positively self-selected rural-to-urban migrants, whereas the remaining rural residents are negatively selected on the basis of health. We disentangle the compositional difference and true health difference by looking into the interaction between hukou and residence-based urban definitions.

3. Materials and methods

3.1 China Family Panel Studies (CFPS)

The data in this study come from the 2014 wave of the China Family Panel Studies (CFPS). We have also run the analysis with the 2012 wave of the data, and the results are qualitatively the same. The 2010 wave of the CFPS survey used different responses options to the SRH question and did not include any anchoring vignettes. The CFPS is a nationally representative longitudinal survey that collects information at the individual, family, and community levels (Xie 2012). The survey has a broad focus and contains information on economic and educational outcomes, migration, family dynamics, and health. An additional advantage of the CFPS dataset is that the three different urban definitions are available. In designing the sampling frame for the CFPS, the Chinese

population was treated as a single entity and therefore was not stratified by any rural–urban definition. The motivation for this was that given China's rapid urbanization, rural–urban administrative definitions are dynamic (Xie and Hu 2014).

The CFPS used multistage probability proportional to size sampling with implicit stratification by administrative units and measures of socioeconomic development: The primary sampling unit was a district (in urban areas) or a county (in rural areas), the second-stage sampling unit was an urban residents' community or a rural village, and the third-stage sampling unit was the household. The 2010 baseline survey successfully interviewed 14,960 households from 635 communities, including 33,600 adults and 8,990 children, located in 25 designated provinces. The approximate response rate was 81%, the majority of the nonresponses being due to a lack of contact. The CFPS's stratified multistage sampling strategy ensures that the sample represents 95% of the total population in China in 2010 (Xie and Hu 2014). The full-scale follow-up surveys in 2012 and 2014 were completed with high follow-up rates of 80.6% and 83.8%, respectively.

The primary sample used in this study is restricted to adults between the ages of 18 and 70. Following recent studies of SRH using the CFPS data (Xu and Xie 2016, 2017), we chose 70 as the upper age limit for both empirical and substantive reasons. Empirically, few respondents (7.7% of the sample after excluding missing data) were above 70 years old as the national average life expectancy was 74.8 years in 2010 (NBS 2012). Substantively, Chinese older adults aged above 70 years have had drastically different life experiences and are less comparable to their younger counterparts with respect to rural–urban gap in health. SRH might not be an appropriate health indicator to study these older adults.

Among a total number of 33,120 adult respondents aged 18 to 70 in the 2014 wave, approximately 27,760 (or 84%) were retained in our final analysis. Missing data are predominantly due to 4,738 (or 14%) adults who did not respond to the anchoring vignettes. As a group, this 14% of respondents had lower SES and poorer health compared to those with valid ratings of the vignettes. The missing data pattern in our sample is consistent with prior research using the 2012 wave of the CFPS data (Xu and Xie 2016, 2017), and our estimates are likely conservative and reflect the lower bound of socioeconomic disparity in health.

In the 2014 wave of the CFPS, SRH was collected by asking respondents the question "How would you rate your health status?" The response categories were poor, fair (this option was not read aloud), good, very good, and excellent. As previously mentioned, the CFPS contains anchoring vignettes. The text of the anchoring vignette used in the CFPS survey can be found in Supplementary Material A. Regarding the 'fair' category, Chinese respondents tend to choose the middle category instead of the extreme responses, regardless of their true opinions/attitudes. During the face-to-face or telephone interviews, the CFPS interviewers read out all the other response categories except 'fair'

on purpose to reduce the likelihood of such reporting behaviors. Nevertheless, respondents could report "fair" (and be recorded so), even though they were not informed of this response category beforehand.

3.2 Statistical analyses

We use the cross-sectional full sample weights to calculate the descriptive statistics. Sampling weights are also used in all the regressions, and the standard errors are adjusted for clustering of the sample by family ID.

To address the first two aims – (1) is there an urban SRH advantage and (2) is it sensitive to how we define urban – we compare the distribution of responses to the SRH question by the different types of urban definitions: administrative, statistical, and based on hukou status. We also look at how the distribution of SRH outcomes varies with the interaction between hukou status and the residence-based definitions of urban.

3.2.1 Models with controls for age and sex

Age and sex are well-known health confounders; therefore, we also analyze whether rural–urban differences in SRH are influenced by the age and sex composition. Given that our research aim is to analyze how rural/urban status is associated with SRH and how that association depends on how urban is measured, we do not control for any mediators (variables such as socioeconomic status and diet that are on the causal pathway between urban treatment and SRH). Including mediators as controls in our regressions can lead to biased estimates of the association between urban status and SRH (Acharya, Blackwell, and Sen 2016).

We first tested whether ordered probit/logit models could be used to control for age and sex. The Brant test indicated that the proportionate odds assumption was violated, therefore using ordered probit/logit models was not appropriate (Long and Freese 2006).

We control for age (age and age squared) and sex by running multinomial logit regressions, which can be used for ordinal outcomes but does not make explicit use of the ordered response categories. A disadvantage of using a multinomial model is that there are multiple coefficients on the treatment variable, one for each of the SRH outcome categories. We present the adjusted predicted probabilities of the SRH outcomes for each of the different rural/urban categories. The average adjusted predicted probabilities (analogous to average marginal effects) means that the values of the other independent variables, age and sex, are left as is.

3.2.2 HOPIT-models correcting for reporting heterogeneity

Hierarchical ordered probit (HOPIT) models use the information in the anchoring vignettes to correct for reporting heterogeneity (Bago d'Uva et al. 2008). The HOPIT models allow the cut-points to vary with covariates. By fitting the HOPIT models to the CFPS data, we can assess patterns of reporting heterogeneity and obtain bias-adjusted estimates of health disparities. We assess parallel versus nonparallel cut-point shifts by estimating two nested models and performing Wald tests. For each covariate, we perform a Wald test in which the null hypothesis is that the same covariate shifts all cut-points in the same direction (upward or downward) by the same magnitude (i.e., parallel shifts). When the null hypothesis is rejected by a Wald test, it indicates that a HOPIT model of nonparallel shifts fits the data better than the one of parallel shifts. To simplify the interpretation of the results we also present the adjusted predictive probabilities.

4. Results

4.1 Sample descriptive statistics

The different urban definitions are correlated but capture different populations. The proportions of our sample considered urban varies based on the definition used. If using the administrative definition and hukou status, 33% and 31%, respectively, of the sample are considered urban. On the other hand, 60% of our sample is considered urban if using the statistical definition (Table 1). Approximately 30% of the individuals considered rural based on the administrative definition (6,299) live in densely populated areas and are considered urban based on the statistical definition.

Table 1: **Descriptive statistics for analytic samples, China CFPS 2014**

	Administrative			Hukou			NBS Statistical		
	Rural	Urban	Diff	Rural	Urban	Diff	Rural	Urban	Diff
Number of obs.	20,245	7,515		20,096	7,664		14,464	13,296	
Obs. As %	67.2	32.8		68.9	31.1		39.8	60.2	
Male (%)	49.5	48.8	0.7	48.7	50.5	1.8	49.0	49.4	0.4
Age	43.2	41.8	1.4	42.7	42.9	0.3	44.4	41.7	2.8
Age groups (%)									
18–30	23.7	24.4	0.7	24.7	22.4	2.3	21.4	25.7	4.3
31–40	17.9	22.5	4.6	18.4	21.7	3.3	16.1	21.6	5.5
41–50	26.0	24.7	1.3	26.3	23.8	2.6	26.2	25.1	1.1
51–60	18.4	18.3	0.1	17.6	20.1	2.6	20.5	17.0	3.5
61–70	14.0	10.1	3.9	13.0	12.0	1.0	15.8	10.7	5.1
Education (%) [n = 25,523]									
No school	28.6	8.7	19.8	28.4	8.1	20.4	33.1	14.8	18.3
Primary school	25.5	12.0	13.5	25.7	10.9	14.7	26.7	17.4	9.3
Junior high school	32.5	32.3	0.3	33.6	29.8	3.8	29.4	34.5	5.1
Senior high school	10.5	25.8	15.3	9.9	27.8	17.9	8.8	20.0	11.2
College	2.9	21.2	18.3	2.4	23.3	20.9	2.1	13.4	11.3
Marital status (%)									
Unmarried	13.8	17.5	3.7	14.2	16.6	2.4	12.9	16.4	3.4
Married/cohabitating	81.8	77.3	4.5	81.2	78.3	2.9	82.4	78.9	3.6
Divorced/widowed	4.5	5.2	0.8	4.5	5.1	0.5	4.6	4.8	0.1
Net family assets [n = 27,215]	336,391	656,170	319,780	333,738	680,170	346,433	268,318	556,189	287,871
Per capita family income [n = 26,445]	12,625	23,647	11,022	12,794	23,829	11,036	11,532	19,437	7,906
Proportion with medical facility in community [n = 24,347]	86.4	85.8	0.6	87.5	83.0	4.5	83.4	88.3	5.0

Note: There are 27,760 observations with nonmissing information for the different urban definitions. Means and percentages are based on sample weights.

The descriptive statistics in Table 1 highlight the different characteristics of the population by urban definition. On average the urban population is younger (the mean age is not statistically different based on hukou status) and more likely to be unmarried than the rural population. Regardless of the urban definition, urban individuals are more educated than their rural counterparts, but the rural–urban disparity in educational attainment is magnified when comparing individuals of different hukou status or based on administrative definition. Access to medical facility is greater for individuals in densely populated communities (statistical urban definition); there is not much difference in access when comparing rural–urban individuals based on administrative definition; however, rural hukou holders were more likely to have a medical facility in their

community compared to urban hukou holders. The last finding is unexpected given the common perception of rural disadvantage in access to health care. Through additional analysis (results not shown), this unusual finding seems to be driven by two factors. First, rural-to-urban migrants were more likely to self-select into urban communities with better socioeconomic conditions, which are not surprisingly also equipped with greater access to medical facilities compared with urban natives. Second, there might be a rural disadvantage in the quality of community medical facilities, despite a rural advantage in access. Many medical facilities in rural communities were small medical clinics. In contrast, although relatively fewer urban communities were equipped with a medical facility, they might be in close proximity to hospitals located within the city boundary.

4.2 Model results

Figure 2 shows the weighted proportion of individuals by urban classification that rate their health as poor, fair, good, very good, or excellent in our sample. The modal response for urban and rural individuals is to rate their health as good; on average more than a third choose this response. It appears that rural respondents, regardless of urban definition used, are more likely (by about five percentage points) to rate their health as poor compared to urbanites (15% administratively rural compared to 10% administratively urban residents, 16% statistical rural versus 11% of statistically urban, and 14% rural hukou status holders versus 10% urban hukou status holders). They are also more likely (by three to six percentage points) to rate their health as excellent or very good compared to urbanites (38% administratively rural compared to 35% administratively urban residents, 38% statistical rural versus 36% of statistically urban, and 39% rural hukou status holders versus 33% urban hukou status holders). If we dichotomize SRH by collapsing the probabilities, as good health is equal to a rating of good health or better, we do find an urban advantage regardless of the urban definition used (see Figure 2; 71% administratively rural compared to 76% administratively urban residents, 70% statistical rural versus 74% of statistically urban, and 71% rural hukou status holders versus 75% urban hukou status holders). Controlling for age and sex does not have a large effect but does reduce the magnitude of the rural–urban disparity in SRH across all the different rural–urban definitions (there is a two to four percentage point difference between rural and urban residents; see Figure 3 for more details). Urban hukou status holders and administrative residents have the largest probability of rating their SRH as good (42%), consequently they have the highest probability of rating their health as good or better. However, residents in statistically urban locations were a little more likely to rate their health as very good or excellent (35%) compared to the other classified categories of urban residents (33% to 34%).

Figure 2: **Rural and urban differences in the probability of describing SRH as poor, fair, good, very good, or excellent by the different Chinese urban definitions in CFPS 2014 dataset**
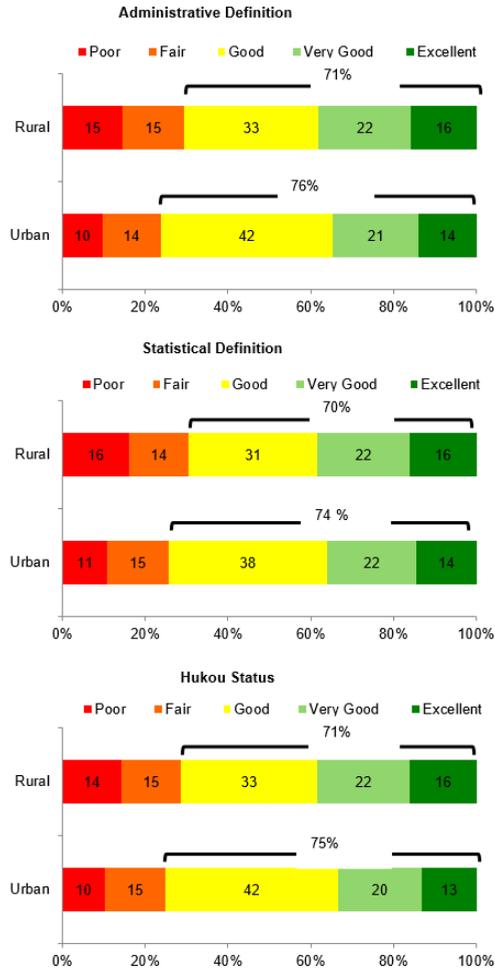
Note: Data labels may not add up to 100 due to rounding. The brackets show the results of collapsing the probabilities of rating SRH as good or better; readers can also choose to collapse at other cut-offs.

Figure 3:   **Adjusted predicted probabilities of different SRH outcomes by urban definition from multinomial logistic regression in CFPS 2014 dataset**
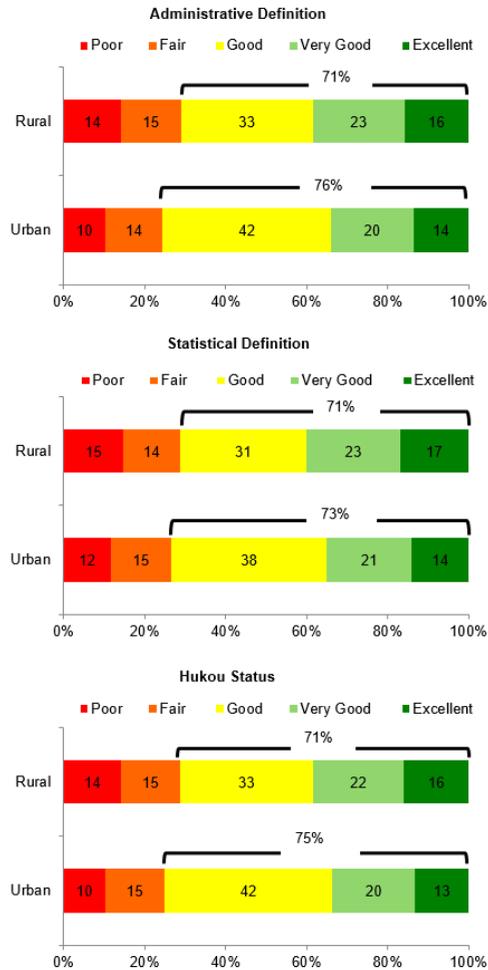
Note: These results control for age and sex. Data labels may not add up to 100 due to rounding. The brackets show the results of collapsing the probabilities of rating SRH as good or better; readers can also choose to collapse at other cut-offs. An accompanying table with 95% confidence intervals can be found in Supplement C.

Next, we turn to the HOPIT model results (Table 2 and Figure 4) to determine whether the rural–urban differences in SRH are real or due to reporting heterogeneity, and if there is reporting heterogeneity, what is the direction of the bias? The slight urban advantage in SRH remains after correcting for reporting heterogeneity. The urban SRH is the largest and only evident when using the statistical definition of rural and urban ($\beta = 0.042$; $p = 0.012$). Figure 4 also illustrates this point; the largest difference in latent health is between statistically rural and urban residents. The urban health advantage is somewhat notable when using the administrative definition ($\beta = 0.034$; $p = 0.064$).

Reporting heterogeneity based on rural–urban status does exist in our dataset; the bias does not result in simple parallel shift in the threshold cut-points. The covariates do not affect all the cut-points by the same magnitudes or direction (Table 2). Regardless of urban classification used, the cut-point between "poor" and "fair" health is lower for urban respondents than for rural respondents. In other words, urban respondents are more likely to choose "fair" over "poor" for a given level of true health somewhere between poor and fair; that bias is most evident for urban hukou holders (threshold 1: 0.174; $p < 0.001$). Urban hukou holders and urban administrative residents also maintain a lower cut-point between "fair" and "good" and thus are also more likely to rate health as "good" instead of "fair" compared to rural individuals for a given level of true health somewhere between fair and good. Regardless of urban classification used, urban respondents hold higher cut-points between "good" and "very good" and between "very good" and "excellent." Therefore, on the higher-end of the latent health spectrum, urban respondents are more likely to rate health as "good" instead of "very good" or "very good" instead of "excellent." Urban individuals in our dataset did not have consistently higher or lower health expectations, the bias in their ratings depends on the level of latent health.

The adjusted predicted probabilities further illustrate the impact of correcting for reporting heterogeneity. Overall the changes are subtle, but we see that the probability of urban individuals rating their health as "fair" or "poor" has increased; the adjustment is the largest for urban hukou status holders (Table 2 and Figure 5). Both rural and urban individuals overestimated their SRH; after correcting for heterogeneity they were less likely to report "very good" or "excellent" health. Nevertheless, rural individuals are still more likely to rate their health as "very good" or "excellent" compared to urban individuals, but the difference is largest when comparing rural and urban individuals based on hukou status and the smallest when comparing rural and urban individuals based on the statistical definition.

Table 2:   HOPIT model results from China CFPS 2014

A.

(N = 23,930)	Vignette-adjusted SRH		Cut-Points								Wald test of parallel shift (df = 3)	
			Poor - Fair		Fair - Good		Good - Very Good		Very Good - Excellent			
	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val
Urban committee (ref: Rural committee)	0.034	0.064	−0.150	0.000	−0.059	0.000	0.118	0.000	0.116	0.000	241.97	0.000
Age	−0.024	0.000	0.034	0.000	0.035	0.000	0.019	0.000	0.001	0.732	94.94	0.000
Age-squared	−0.000	0.093	−0.000	0.000	−0.000	0.000	−0.000	0.000	−0.000	0.079	63.56	0.000
Male (ref: female)	0.283	0.000	0.036	0.003	0.020	0.077	0.050	0.000	0.032	0.014	9.72	0.021
Constant	0.443	0.000	−2.489	0.000	−1.947	0.000	−0.633	0.000	0.599	0.000		
Sigma	1.011	0.000										

B.

(N = 23,930)	Vignette-adjusted SRH		Cut-Points								Wald test of parallel shift (df = 3)	
			Poor - Fair		Fair - Good		Good - Very Good		Very Good - Excellent			
	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val
Urban hukou (ref: Rural hukou)	0.027	0.141	−0.174	0.000	−0.076	0.000	0.134	0.000	0.138	0.000	341.93	0.000
Age	−0.023	0.000	0.034	0.000	0.035	0.000	0.020	0.000	0.001	0.653	92.84	0.000
Age-squared	−0.000	0.085	−0.000	0.000	−0.000	0.000	−0.000	0.000	−0.000	0.053	60.44	0.000
Male (ref: female)	0.282	0.000	0.039	0.001	0.021	0.060	0.048	0.000	0.030	0.021	9.25	0.026
Constant	0.444	0.000	−2.486	0.000	−1.944	0.000	−0.639	0.000	0.592	0.000		
Sigma	1.011	0.000										

C.

(N = 23,930)	Vignette-adjusted SRH		Cut-Points								Wald test of parallel shift (df = 3)	
			Poor - Fair		Fair - Good		Good - Very Good		Very Good - Excellent			
	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val
NBS-defined Urban (ref: Rural)	0.042	0.012	−0.098	0.000	0.000	0.994	0.139	0.000	0.113	0.000	236.98	0.000
Age	−0.024	0.000	0.033	0.000	0.035	0.000	0.020	0.000	0.001	0.648	90.89	0.000
Age-squared	−0.000	0.094	−0.000	0.000	−0.000	0.000	−0.000	0.000	−0.000	0.063	60.44	0.000
Male (ref: female)	0.283	0.000	0.036	0.003	0.022	0.056	0.050	0.000	0.032	0.014	8.67	0.034
Constant	0.432	0.000	−2.472	0.000	−1.954	0.000	−0.678	0.000	0.569	0.000		
Sigma	1.012	0.000										

Figure 4: **Predicted cut-points by rural–urban status from the HOPIT model, using CFPS 2014 dataset. Higher values on the latent health scale indicate better health**
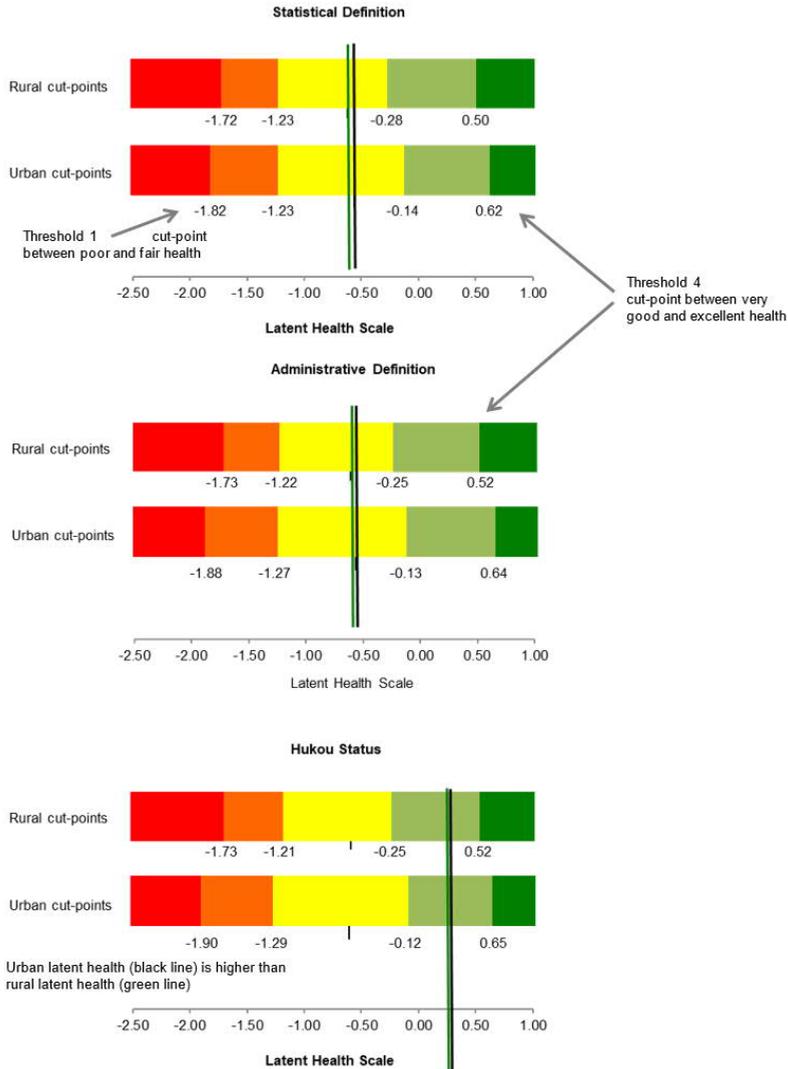
Figure 5:   **Average adjusted predicted probabilities of SRH outcomes from the HOPIT models, which correct for reporting heterogeneity. Data labels may not add up to 100 due to rounding**
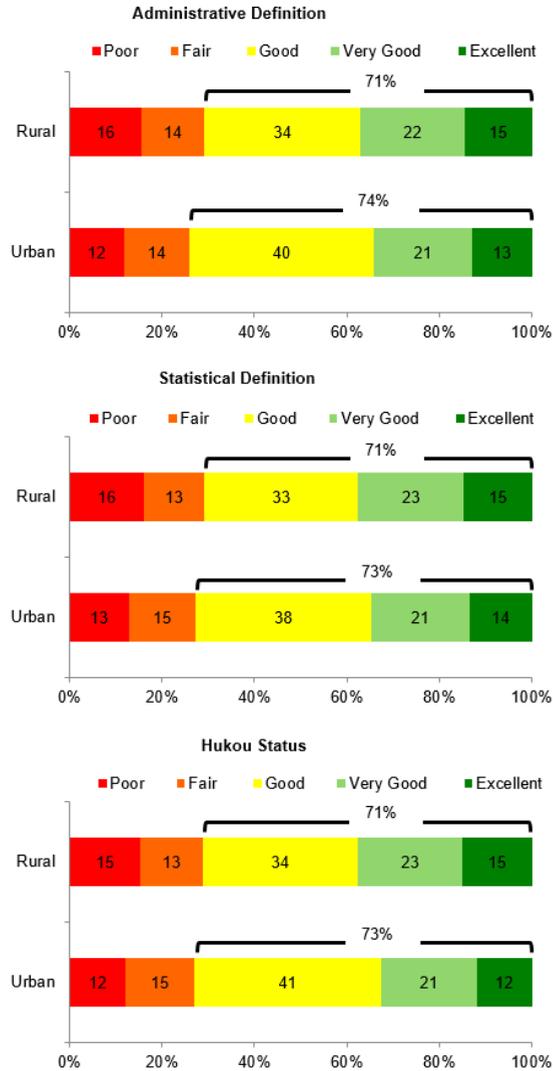
4.3 Interaction between hukou and location based urban definitions

In this section, we look at the interaction between hukou status and location-based urban definitions to disaggregate the impact of residing in different types of urban locations on the SRH status of nonmigrants and migrants.

Rural hukou holders in statistically (NBS) urban areas are not necessarily all migrants since some administratively rural areas are considered urban by the NBS (recall 30% of individuals that are administratively rural live in NBS urban areas). After controlling for age and sex, urban nonmigrants (urban hukou and urban NBS) were most likely to report their health as good or better (75%) followed by urban hukou holders in rural areas (73%), then rural-to-urban migrants (rural hukou and urban NBS, 72%), and rural nonmigrants (rural hukou and rural NBS, 71%)) (Figure A-1, Panel A). After controlling for reporting heterogeneity, the rankings change. There is very little difference in the SRH of urban nonmigrants and rural-to-urban migrants (rural hukou and urban NBS). Residents of dense urban areas regardless of hukou status had higher SRH than those residing in less dense areas (Table 3, Panel A). Therefore, part of the NBS urban advantage in SRH is due to migrant advantage, but it is also due to the fact that residents of dense areas have higher SRH.

Figure A-1, Panel B, shows that after controlling for age and sex composition, rural migrants to urban administrative areas (rural hukou and administrative urban) were the most likely to rate SRH as good or better (76%), followed by urban nonmigrants (75%), then urban hukou holders in rural administrative areas (74%). Rural nonmigrants had the worse SRH (71%). After controlling for reporting heterogeneity, rural migrants to urban administrative areas still have the highest latent health (Table 3, Panel B). This is not surprising since they are a positively selected group. Urban hukou holders in urban administrative (urban nonmigrants) were the most likely to overestimate SRH if in poor or fair health; consequently urban nonmigrants did not have higher SRH compared to rural nonmigrants after adjusting for reporting heterogeneity. Therefore, the majority of the administratively urban SRH advantage is due to migrant advantage.

Table 3: **HOPIT model results for interaction of hukou status and residential based urban definitions**

A.

| | Vignette-adjusted SRH | | Cut-Points | | | | | | | | Wald test of parallel shift (df = 3) | |
| | | | Poor - Fair | | Fair - Good | | Good - Very Good | | Very Good - Excellent | | | |
	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val
Rural hukou & Rural NBS (ref; n=11,690 or 48.9%)												
Rural hukou & Urban NBS (n=5,548 or 23.2%)	0.038	0.068	−0.026	0.153	0.057	0.001	0.101	0.000	0.052	0.009	57.13	0.000
Urban hukou & Rural NBS (n=885 or 3.7%)	0.005	0.900	−0.129	0.001	−0.031	0.382	0.049	0.185	0.006	0.899	22.45	0.000
Urban hukou & Urban NBS (n=5,807 or 24.3%)	0.043	0.034	−0.191	0.000	−0.062	0.000	0.185	0.000	0.179	0.000	392.51	0.000
Age	−0.024	0.000	0.034	0.000	0.035	0.000	0.020	0.000	0.001	0.750	95.92	0.000
Age-squared	−0.000	0.092	−0.000	0.000	−0.000	0.000	−0.000	0.000	−0.000	0.077	63.08	0.000
Male (ref: female)	0.283	0.000	0.038	0.001	0.022	0.051	0.049	0.000	0.032	0.014	8.83	0.032
Constant	0.431	0.000	−2.481	0.000	−1.969	0.000	−0.676	0.000	0.580	0.000		
Sigma	1.011	0.000										

B.

| | Vignette-adjusted SRH | | Cut-Points | | | | | | | | Wald test of parallel shift (df = 3) | |
| | | | Poor - Fair | | Fair - Good | | Good - Very Good | | Very Good - Excellent | | | |
	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val	coef	p-val
Rural hukou & Rural admin (ref; n=15,907 or 66.5%)												
Rural hukou & Urban admin (n=1,331 or 5.6%)	0.089	0.007	−0.046	0.110	0.024	0.389	0.071	0.013	0.080	0.013	16.49	0.001
Urban hukou & Rural admin (n=1,545 or 6.5%)	0.068	0.033	−0.125	0.000	−0.036	0.184	0.121	0.000	0.146	0.000	71.25	0.000
Urban hukou & Urban admin (n=5,147 or 21.5%)	0.022	0.285	−0.194	0.000	−0.087	0.000	0.145	0.000	0.143	0.000	319.16	0.000
Age	−0.023	0.000	0.034	0.000	0.035	0.000	0.020	0.000	0.002	0.628	92.81	0.000
Age-squared	−0.000	0.075	−0.000	0.000	−0.000	0.000	−0.000	0.000	−0.000	0.053	60.76	0.000
Male (ref: female)	0.282	0.000	0.038	0.001	0.021	0.066	0.049	0.000	0.031	0.020	9.5	0.023
Constant	0.428	0.000	−2.485	0.000	−1.951	0.000	−0.647	0.000	0.581	0.000		
Sigma	1.011	0.000										

5. Discussion

As in many other developing countries, the rural–urban divide is one of the most fundamental socioeconomic and demographic markers and a major driving force behind inequalities in China. Rapid economic reforms since the late 1970s have arguably benefited rural and urban populations to different extents (Kanbur and Zhang 1999; Xie and Hannum 1996; Xie and Zhou 2014). Given the strong association between socioeconomic status and health, the resulting widening socioeconomic inequalities between the rural and urban populations would translate into rising rural–urban health disparities. For these reasons, numerous studies have sought to document the pattern and trend of rural–urban health disparities in China. However, previous research on this topic tends to overlook the measurement issue of varying urban definitions. We were motivated to accurately measure rural–urban disparity in SRH in China and examine if choice of urban definition played a role. We expanded previous research that usually adopted one urban definition by examining two community-level definitions and one individual-level definition.

Surprisingly, we found that before controlling for reporting heterogeneity, the choice of urban definition did not alter the estimated rural–urban difference in SRH. The distributions of SRH across the different rural–urban classifications are very similar (Figures 2 and 3). Nevertheless, to better compare findings across studies, researchers that use Chinese data should clearly state which urban definition they are using when studying rural–urban disparities. Although rural–urban hukou status continues to convey unequal civil rights and life opportunities to date (Li, Gu, and Zhang 2015), our results suggest that its association with SRH is not considerably stronger compared to other urban definitions. If one has to make a single choice for studying rural–urban disparities in SRH in China, the statistical definition of urban seems to be a more appropriate option. Driven by growth of the nonagricultural population and expansion of urban land use, the statistical definition is more dynamic and can more easily capture the urbanization process. It is also annually updated by China's National Bureau of Statistics. In contrast, as mentioned in the theoretical background, both the hukou-based and administrative definitions are much more time-static and lagged behind in capturing fast urbanization of a community's socioeconomic and built environments. Future work might want to move away from a dichotomous definition of rural–urban status, which may miss much of the variation within the categories. Nevertheless, we have considered only SRH in this study, and it is possible that choice of urban definition might have a larger impact on the estimation of rural–urban health disparities when looking at different outcomes (Zhu and Osterle 2017).

Building on the literature that uses SRH as the main health indicator, we also addressed another important measurement issue: reporting heterogeneity. As several

studies have suggested, empirical estimates of rural–urban difference in SRH result from a combination of true health disparity and reporting heterogeneity in SRH due to different health expectations held by rural and urban Chinese (Bago d'Uva et al. 2008; van De Poel, Van Doorslaer, and O'Donnell 2012; Whyte and Sun 2010; Zimmer and Kwong 2004). Most of these studies have conjectured that given their higher social status, the urban Chinese are likely to hold a higher health expectation than their rural peers and thus underreport their true health status, resulting in an underestimated urban health advantage. Using anchoring vignette data, we contributed to this literature by testing whether there is any systematic rural–urban difference in health expectation and self-rating and, if so, to what extent such reporting heterogeneity affects the estimate of true rural–urban health disparity.

Consistent with an earlier study (Bago d'Uva et al. 2008), we found a systematic difference between rural and urban Chinese in the ways in which they self-assess health status. Using regional data collected in three Chinese provinces in 2000–2001, Bago d'Uva et al. (2008) adopt the statistical urban definition and report cut-point shifts by rural–urban residence in five out of six self-reported health domains (mobility, cognition, pain, self-care, and usual activities, but not affect). Unfortunately, they did not present or compare the estimates of rural–urban difference in health before and after purging out reporting heterogeneity. We found persistent reporting heterogeneity by rural–urban status, regardless of the choice of urban definition. Contrary to our expectation, however, the urban respondents do not always hold a higher standard for self-assessment of health. Instead, the cut-point shifts in their self-ratings depend on the level of latent health. For a given level of true poor health, urban respondents would be more likely than rural respondents to report fair (versus poor) or good (versus fair) health, thereby inflating the rural–urban health gap. For a given level of true good health, however, urban respondents would be less likely to report very good (versus good) health or excellent (versus very good) health, thereby deflating the rural–urban health gap. One possible explanation is that urban respondents tend to invoke health optimism – higher-status people believe that their affluence confers well-being and hence systematically boost their self-ratings of health (Ferraro 1980). As a result, they tend to consider themselves in fair health – the downward shift of the cut-point between poor and fair – when they are in poor health. But when urban respondents are in very good or excellent health, they tend to compare themselves with their peers and adopt a higher standard for self-rating – the upward shift of the cut-points between good and very good and between very good and excellent – and thus underrate their true health status.

Our results have implications for future studies of SRH. There is always the possibility that reporting heterogeneity is biasing SRH results and it is very difficult to know the magnitude or direction of the bias. In recent studies of the CFPS data, Xu and Xie (2016 and 2017) find systematic variation by sociodemographic characteristics in

respondents' reporting behaviors. However, the impact of reporting heterogeneity on biasing the estimate of health disparities was most evident with respect to respondents' educational attainment. Our paper also highlights the fact that researchers should be cautious in how they dichotomize SRH; we found that reporting heterogeneity could deflate or inflate SRH based on the cutoffs we chose.

The statistical definition of urban may be better able to capture the built environment, which might explain why we consistently find an urban SRH advantage for residents in densely populated areas. This is in contrast with some of the existing literature, such as a cross-sectional study finding that individuals living in more populous neighborhoods in Brazil had lower rating of SRH compared to individuals living in less populous neighborhoods (Cremonese et al. 2010). In this literature, higher population density is typically associated with factors such as overcrowding and higher levels of pollution, which can have negative effects on health and SRH. On the other hand, we might expect a positive association if the provision of many services is density based (Hanlon et al. 2012). Residents in densely populated areas might have greater SRH because of increased access to density dependent services (see Table 1, proportion with medical facility in community). However, it is not clear if access to services might lead to higher SRH – Subramanian et al. (2006) does not find an association between neighborhood service provision and good SRH in the elderly. Knowledge of available services might also spread more easily in densely population areas (Rank and Hirschl 1993).

Putting our different analyses together, can we posit an urban advantage or disadvantage in SRH in contemporary China? Yes, there is a slight urban advantage in SRH, but after controlling for reporting heterogeneity it is most pronounced and evident only when using the statistical definition of urban. Part of the urban advantage in SRH comes from the positively selected rural-to-urban migrants, which have the highest SRH. The small and relatively privileged population of urban hukou holders in rural areas also had high levels of SRH. Our results differ from some of the previously cited literature, which finds no advantage or an urban disadvantage in SRH in China. The key difference is that these papers controlling for other factors, which we believe are mediators (i.e., may be causally affected by rural–urban status), might lead to bias (Angrist and Pischke 2009). Second, by adding additional 'controls' these authors are trying to identify the independent effect of urban residence on SRH, while we asked how rural–urban residence is associated with SRH. If we control for socioeconomic characteristics (education, family assets, and per capita income) and do not conduct proper mediation analysis, we find that the urban advantage disappears and becomes a slight rural advantage (1 to 2 percentage points difference in the likelihood of rating health as good or better).

An important methodological implication of our findings is that urbanness is multidimensional and can be defined and measured in different ways. The research into

rural–urban disparities in health (or other domains) should be mindful and explicit about the choice of the urban definition. Such a choice depends on what aspect of health is under investigation. In the Chinese context, for example, distribution of health care may be heavily constrained along administrative and institutional (e.g., hukou) boundaries, whereas environmental health may depend more upon demographic conditions. Future research should make a theoretically grounded choice of the urban definition in order to draw meaningful inference.

This study has several limitations. First, the validity of the anchoring vignette analysis hinges on certain modeling assumptions. For example, we assume that the vignette equivalence assumption holds when rural and urban respondents might not interpret the same vignette in the same way. We also assume that the response consistency assumption holds when a given respondent, rural or urban, might not apply the same standard for ratings of the vignettes as that for self-ratings of health. If this is the case the anchoring vignettes might not be correcting for reporting heterogeneity. Formally testing these modeling assumptions requires additional data such as a good proxy of true health status, which is beyond the scope of this study. Nevertheless, a recent study has shown that the CFPS anchoring vignettes are effective in adjusting for reporting heterogeneity by sociodemographic characteristics (Xu and Xie 2016).

Second, the current analysis cannot explain why there is an urban advantage in SRH. Future research will use mediation analysis to identify mediators responsible for the urban advantage in SRH. We plan to test mediators in the following domains: socioeconomic status, lifestyle, access to health care, and environment. We also plan to compare the results of the mediation analysis with regressions that control for mediators in standard regression to see if the latter, which is widely used in the literature, produces biased results (Acharya et al. 2016).

6. Acknowledgement

References

Acharya, A., Blackwell, M., and Sen, M. (2016). Explaining causal findings without bias: detecting and assessing direct effects. *The American Political Science Review* 110(3): 512–529. doi:10.1017/S0003055416000216.

Angrist, J.D. and Pischke, J-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press. doi:10.1515/97814008 29828.

Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., and O'Donnell, O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics* 17(3): 351–375. doi:10.1002/hec.1269.

Benjamins, M.R., Hummer, R.A., Eberstein, I.W., and Nam, C.B. (2004). Self-reported health and adult mortality risk: An analysis of cause-specific mortality. *Social Science and Medicine* 59(6): 1297–1306. doi:10.1016/j.socscimed.2003.01.001.

Campbell, C. (1997). Public health efforts in China before 1949 and their effects on mortality: The case of Beijing. *Social Science History* 21(2): 179–218. doi:10.1017/S0145553200017703.

Chan, K.W. (2007). Misconceptions and complexities in the study of China's cities: Definitions, statistics, and implications. *Eurasian Geography and Economics* 48(4): 383–412. doi:10.2747/1538-7216.48.4.383.

Chan, K.W. (2013). China: Internal migration. In: *The Encyclopedia of Global Human Migration*: Blackwell Publishing Ltd. 67–72. doi:10.1002/9781444351071. wbeghm124.

Chen, J. (2011). Internal migration and health: Re-examining the healthy migrant phenomenon in China. *Social Science and Medicine* 72(8): 1294–1301. doi:10.1016/j.socscimed.2011.02.016.

Chen, J., Chen, S., and Landry, P.F. (2013). Migration, environmental hazards, and health outcomes in China. *Social Science and Medicine* 80: 85–95. doi:10.1016/ j.socscimed.2012.12.002.

Cremonese, C., Backes, V., Olinto, M.T.A., Dias-da-Costa, J.S., and Pattussi, M.P. (2010). Neighborhood sociodemographic and environmental contexts and self-rated health among Brazilian adults: A multilevel study. *Cadernos de Saude Publica* 26(12): 2368–2378. doi:10.1590/S0102-311X2010001200015.

Do, D.P, Frank, R., and Iceland, J. (2017). Black-white metropolitan segregation and self-rated health: Investigating the role of neighborhood poverty. *Social Science and Medicine* 187: 85–92. doi:10.1016/j.socscimed.2017.06.010.

Dorélien, A., Balk, D., and Todd, M. (2013). What is urban? Comparing a satellite view with the demographic and health surveys. *Population and Development Review* 39(3): 413–439. doi:10.1111/j.1728-4457.2013.00610.x.

Dowd, J.B. and Todd, M. (2011). Does self-reported health bias the measurement of health inequalities in U.S. adults? Evidence using anchoring vignettes from the health and retirement study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 66B(4): 478–489. doi:10.1093/geronb/gbr050.

Ferraro, K.F. (1980). Self-ratings of health among the old and the old-old. *Journal of Health and Social Behavior* 21(4): 377–383. doi:10.2307/2136414.

Grol-Prokopczyk, H., Freese, J., and Hauser, R.M. (2011). Using anchoring vignettes to assess group differences in general self-rated health. *Journal of Health and Social Behavior* 52(2): 246–261. doi:10.1177/0022146510396713.

Hanlon, M., Burstein, R., Masters, S.H., and Zhang, R. (2012). Exploring the relationship between population density and maternal health coverage. *BMC Health Services Research* 12: 416–422. doi:10.1186/1472-6963-12-416.

Idler, E.L. and Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior* 38(1): 21–27. doi:10.2307/2955359.

Kanbur, R. and Zhang, X. (1999). Which regional inequality? The evolution of rural–urban and inland–coastal inequality in China from 1983 to 1995. *Journal of Comparative Economics* 27(4): 686–701. doi:10.1006/jcec.1999.1612.

King, G., Murray, C.J.L., Salomon, J.A., and Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review* 98(1): 191–207. doi:10.1017/S000305540400108X.

Li, J., Gu, Y., and Zhang, C. (2015). Hukou-based stratification in urban China's segmented economy. *Chinese Sociological Review* 47(2): 154–176. doi:10.1080/21620555.2014.990326.

Long, J.S. and Freese, J. (2006). Regression models for categorical and limited dependent variables using Stata, second edition. College Station: Stata Press.

Luo, Y. and Wen, M., (2002). Can we afford better health? A study of the health differentials in China. *Health* 6(4): 471–500. doi:10.1177/136345930200600404.

Ma, L.J.C. and Cui, G. (1987). Administrative changes and urban population in China. *Annals of the Association of American Geographers* 77(3): 373–395. doi:10.1111/j.1467-8306.1987.tb00165.x.

NBS (2012). The average life expectancy at birth in China has reached 74.83 years. Beijing: National Bureau of Statistics of China.

Pan, Y., Lin, L.H., and Peng, Z.H. (2012). *Big construction site – living picture of the construction's migrant workers.* Beijing: Peking University Press.

Pei, X. and Rodriguez, E. (2006). Provincial income inequality and self-reported health status in China during 1991–7. *Journal of Epidemiology and Community Health* 60(12): 1065–1069. doi:10.1136/jech.2005.043539.

Qin, B. and Zhang, Y. (2014). Note on urbanization in China: Urban definitions and census data. *China Economic Review* 30*: 495–502. doi:10.1016/j.chieco.2014.07.008.

Qiu, P., Caine, E., Yang, Y., Chen, Q., Li, J., and Ma, X. (2011). Depression and associated factors in internal migrant workers in China. *Journal of Affective Disorders* 134(1–3): 198–207. doi:10.1016/j.jad.2011.05.043.

Rank, M.R. and Hirschl, T.A. (1993). The link between population density and welfare participation. *Demography* 30(4): 607–622. doi:10.2307/2061809.

Schnittker, J. (2005). When mental health becomes health: Age and the shifting meaning of self-evaluations of general health. *The Milbank Quarterly* 83(3): 397–423. doi:10.1111/j.1468-0009.2005.00407.x.

Shen, J. (2006). Understanding dual-track urbanisation in post-reform China: Conceptual framework and empirical analysis. *Population, Space and Place* 12(6): 497–516. doi:10.1002/psp.426.

Singh-Manoux, A., Pekka, M., Ferrie, J., Zins, M., Marmot, M., and Goldberg, M. (2006). What does self rated health measure? Results from British Whitehall II and French Gazel cohort studies? *Journal of Epidemiology and Community Health* 60(4): 364–372. doi:10.1136/jech.2005.039883.

Song, Y. and Sun, W. (2016). Health consequences of rural-to-urban migration: Evidence from panel data in China. *Health Economics* 25(10): 1252–1267. doi:10.1002/hec.3212.

Subramanian, S.V., Kubzansky, L.D., Berkman, L., Fay, M., and Kawachi, I. (2006). Neighborhood effects on the self-rated health of elders: Uncovering the relative importance of structural and service-related neighborhood environments. *The*

Journals of Gerontology Series B: Psychological Sciences and Social Sciences 61(3): S153–S160. doi:10.1093/geronb/61.3.S153.

Tong, Y. and Piotrowski, M. (2012). Migration and health selectivity in the context of internal migration in China, 1997–2009. *Population Research and Policy Review* 31(4): 497–543. doi:10.1007/s11113-012-9240-y.

United Nations, Department of Economic and Social Affairs, Population Division (2015). World urbanization prospects: The 2014 revision. New York, NY: United Nations. (No. ST/ESA/SER.A/366).

Utzinger, J. and Keiser, J. (2006). Urbanization and tropical health – then and now. *Annals of Tropical Medicine and Parasitology* 100(5–6): 517–533. doi:10.1179/136485906X97372.

Van de Poel, E., O'Donnell, O., and Van Doorslaer, E. (2009). Urbanization and the spread of diseases of affluence in China. *Economics and Human Biology* 7(2): 200–216. doi:10.1016/j.ehb.2009.05.004.

Van de Poel, E., Van Doorslaer, E., and O'Donnell, O. (2012). Measurement of inequity in health care with heterogeneous response of use to need. *Journal of Health Economics* 31(4): 676–689. doi:10.1016/j.jhealeco.2012.05.005.

Whyte, M.K. and Sun, Z. (2010). The impact of China's market reforms on the health of Chinese citizens: Examining two puzzles. *China: An International Journal* 8: 1–32. doi:10.1142/S0219747210000038.

Woods, R. (2003). Urban-rural mortality differentials: An unresolved debate. *Population and Development Review* 29(1): 29–46. doi:10.1111/j.1728-4457.2003.00029.x.

Xie, Y. (2012). *The user's manual of the China family panel studies.* Beijing: Institute of Social Science Survey, Peking University.

Xie, Y. and Hannum, E. (1996). Regional variation in earning inequality in reform-era urban China. *The American Journal of Sociology* 101(4): 950–992. doi:10.1086/230785.

Xie, Y. and Hu, J. (2014). An introduction to the China family panel studies (CFPS). *Chinese Sociological Review* 47: 3–29.

Xie, Y. and Zhou, X. (2014). Income inequality in today's China. *Proceedings of the National Academy of Sciences* 111(19): 6928–6933. doi:10.1073/pnas.1403158111.

Xu, H. and Short, S.E. (2011). Health insurance coverage rates in 9 provinces in China doubled from 1997 to 2006, with a dramatic rural upswing. *Health Affairs* 30(12): 2419–2426. doi:10.1377/hlthaff.2010.1291.

Xu, H. and Xie, Y. (2016). Assessing the effectiveness of anchoring vignettes in bias reduction for socioeconomic disparities in self-rated health among Chinese adults. *Sociological Methodology* 46(1): 84–120. doi:10.1177/0081175015599808.

Xu, H. and Xie, Y. (2017). Socioeconomic inequalities in health in China: A reassessment with data from the 2010–2012 China family panel studies. *Social Indicators Research* 132(1): 219–239. doi:10.1007/s11205-016-1244-2.

Zajacova, A., Huzurbazar, S., and Todd, M. (2017). Gender and the structure of self-rated health across the adult life span. *Social Science and Medicine* 187: 58–66. doi:10.1016/j.socscimed.2017.06.019.

Zhang, X. and Kanbur, R. (2005). Spatial inequality in education and health care in China. *China Economic Review* 16(2): 189–204. doi:10.1016/j.chieco.2005.02.002.

Zhu, Y. and Österle, A. (2017). Rural-urban disparities in unmet long-term care needs in China: The role of the hukou status. *Social Science and Medicine* 191: 30–37. doi:10.1016/j.socscimed.2017.08.025.

Zimmer, Z. and Amornsirisomboon, P. (2001). Socioeconomic status and health among older adults in Thailand: An examination using multiple indicators. *Social Science and Medicine* 52(8): 1297–1311. doi:10.1016/S0277-9536(00)00232-X.

Zimmer, Z. and Kwong, J. (2004). Socioeconomic status and health among older adults in rural and urban China. *Journal of Aging and Health* 16(1): 44–70. doi:10.1177/0898264303260440.

Zimmer, Z., Natividad, J., Lin, H-S., and Chayovan, N. (2000). A Cross-national examination of the determinants of self-assessed health. *Journal of Health and Social Behavior* 41(4): 465–481. doi:10.2307/2676298.

Appendix

Figure A-1: **Adjusted predicted probability for multinomial logit model that controls for age and sex. Data labels may not add up to 100 due to rounding**

A. Interaction between *hukou* status and statistical definition

B. Interaction between *hukou* status and administrative definition