

# **Construction and application of a revised satisfaction index model for Chinese urban and rural residents basic medical insurance**

3

4

5 Wenwei Cheng<sup>1¶</sup>, Jin Cheng<sup>1¶</sup>, Xiaofang Liu<sup>1</sup>, Yanyan Wu<sup>1</sup>, Weichu Sun<sup>2</sup>, Xiaofang Yan<sup>1</sup>, Lai  
6 Peng<sup>1</sup>, Xiaoli Liu<sup>1</sup>, Qi Wang<sup>1</sup>, Mingming Luo<sup>1</sup>, Tingting Sha<sup>1</sup>, Jingcheng Shi<sup>1\*</sup>, Fang Yang<sup>1\*</sup>

7

8

9

<sup>10</sup> Department of Epidemiology and Medical Statistics, Central South University, Changsha,  
<sup>11</sup> Hunan, China

12 <sup>2</sup>Department of Orthopedics, University of South China, Hengyang, Hunan , China

13

14

15 \*Corresponding author

16 E-mail: shijch@csu.edu.cn(JS), yangfang2010@csu.edu.cn(FY)

17

18

<sup>19</sup>These Authors contributed equally to this work.

20

## 21      **Abstract**

22      **Background:** Quality is the most important factor of satisfaction. However, the existing  
23      URRBMI index model lacks the decomposition of the connotation of perceived quality and  
24      cannot provide a reference for quality improvement and satisfaction promotion.

25      **Objective:** This study aims to construct a Satisfaction Index Model for Chinese Urban and  
26      Rural Residents Basic Medical Insurance (SIM-URRBMI) with accurate and detailed  
27      measurement of perceived quality and give a feasible and scientific suggestion for URRBMI  
28      or insurance for other countries in the world.

29      **Methods:** Based on the theoretical framework of The American Customer Satisfaction Index  
30      (ACSI), the connotation of perceived quality was refined by literature review and expert  
31      consultation to form a pool of alternative measurement variables. A three-stage randomized  
32      stratified cluster sampling was adopted to select the main decision makers for pupils'  
33      URRBMI in 8 primary schools from Changsha City. Both Classic Test Theory (CTT) and  
34      Item Response Theory (IRT) were used for selection of the measurement variables. The  
35      model's reliability and validity were tested using partial least square (PLS) related methods.

36      **Results:** A total of 1909 respondents who had insurance for their children were investigated  
37      with the initial questionnaire. The revised SIM-URRBMI consists of 11 latent variables and  
38      28 measurement variables with good reliability and validity. Among the three explanatory  
39      variables of public satisfaction, perceived quality had the largest total effect (0.737). The  
40      variable with greatest effect among the five first-order latent variables on perceived quality  
41      was quality of the medical insurance policy (0.472).

42      **Conclusions:** The revised SIM-URRBMI consists of 11 latent variables and 28 measurement

43 variables with good reliability and validity. It provides accurate assessment of perceived  
44 quality, which will greatly help performance improvement. Perceived quality is crucial to  
45 public satisfaction, especially, the most important aspects are policies regarding medical  
46 insurance reimbursement (basic coverage scope, coinsurance, deductible).

47 **Keywords:** Satisfaction index model; Basic Medical Insurance for Urban and Rural Residents;

48 PLS-SEM; reliability; validity

## 49      **Introduction**

50      According to the 2015 World Health Organization (WHO) report, 6% of people in low-  
51      and middle-income countries are forced into extreme poverty because of medical  
52      expendituresThe Chinese government has been trying to establish a multilevel medical  
53      security system that is mainly based on the social basic medical insurance (BMI) schemes so  
54      as to protect the finances of individuals and households affected by illness and injury to some  
55      extent. In order to improve equity, sustainability, and efficiency, China has taken an initial  
56      step to integrate new cooperative medical scheme (NCMS) and urban resident-based basic  
57      medical insurance (URBMI) into the urban and rural residents' basic medical insurance  
58      scheme (URRBMI) [1-2].

59      Voluntary enrollment is a principle of URRBMI in China [3]. The enrollee's satisfaction  
60      and loyalty are crucial for the scheme's effective implementation and sustained development.  
61      Moreover, analyzing the enrollee's satisfaction with their health insurance and its influencing  
62      factors can help to shed light on current and future policy decision making [0, 4-4]. ACSI is  
63      one of the most widely used and representative quantitative measurement scales that shows  
64      the cause-and-effect relationship linking the casual variables of customer satisfaction to their  
65      consequent effect variable (customer loyalty) [6]. The model construction theory of ACSI  
66      provides an important reference for assessing the attitude and major improvement direction of  
67      the URRBMI from the perspective of insured participants. Using the ACSI as a basis and  
68      considering the Chinese social and cultural backgrounds, Peng et al. constructed a satisfaction  
69      index model for URRBMI that demonstrated good reliability and validity for satisfaction  
70      measurement[7-8]. There has been sufficient literature supporting the notion that perceived

71 quality may be the most worthwhile topic in service area satisfaction research, and  
72 satisfaction is mainly determined by perceived quality [6, 9-10]. However, Peng's model only  
73 adopted a few indicators to assess perceived quality, such as the overall quality assessment  
74 question, "how much help does the URRBMI bring to your child?". The lack of  
75 comprehensive and detailed aspects of URRBMI quality measurement greatly limit its further  
76 use in health reform [7-8].

77 In addition, since significant variation in economic and cultural backgrounds,  
78 accessibility of health services, and reimbursement policies existed between rural and urban  
79 residents before the URRBMI integration, it is necessary to consider the modeling  
80 heterogeneity from a theoretical and practical perspective to learn about the differences  
81 between groups of the insured (rural VS urban) [11].

82 The aims of this study are to revise and validate the SIM-URRBMI and identify the main  
83 quality aspects needed to be improved in order to provide a valid, reliable, and practical  
84 satisfaction measurement tool for URRBMI. In addition, using heterogeneity analysis, we aim  
85 to explore the differences in satisfaction mechanism between rural and urban participants.

## 86    Materials and Methods

### 87    Participants and Sampling

88       This study was conducted in Changsha, the capital city of the Chinese Province of Hunan,  
89       which completed the integration of NCMS and URBMI in 2011. A three-stage randomized  
90       stratified cluster sampling method was used to ensure a representative sample. The first stage  
91       was to randomly select two districts/counties from the nine districts/counties of Changsha  
92       City. The second stage was to randomly select two primary schools for each selected  
93       district/county. Finally, for each selected school, six classes (one class from each grade) were  
94       randomly selected, and questionnaires were sent home; the main family decision maker  
95       responsible for the pupils' URRBMI was instructed to fill out the questionnaires. The  
96       inclusion criteria were: a) had insurance for their children in 2017-2018; b) were inhabitants  
97       of a city or village in Changsha; c) were the main decision maker for pupils' URRBMI in their  
98       family. Both the students in the selected classes and the main family decision makers  
99       responsible for the pupils' URRBMI were subjects of investigation.

100      This study used a graded response model (GRM) to identify the measurement variables.  
101     In order to make the estimated parameters more accurate, a sample size of at least 500 cases is  
102     suitable. The selected samples of measurement variables included in this paper are 574 cases,  
103     which meet the requirements[13]. The sample size for model evaluation was calculated by the  
104     following formula,  $n \geq 50r^2 - 450r + 1100$  ( $r$  is the ratio of measurement variables to latent  
105     variables), which was based on the structural equation model research conducted by Westland  
106     et al. A minimum sample size of 216 was estimated. However, considering the need for  
107     subgroup analysis, a sample size of 1335 was obtained for model evaluation, which met the

108 estimated requirements [13-14].

## 109    **Ethical approval**

110    The study protocol was reviewed and approved by the Medical Ethics Committee at Xiang  
111    -ya School of Public Health, Central South University. Written informed consent was obtai  
112    ned for each participant who was the medical insurance decision maker of the family.

## 113    **Model construction**

### 114    **Phase 1**

115       The programmed decision process, a method to simultaneously develop a scale by  
116       nominal group (including pupils and their family members, local medical insurance  
117       management and service officers, experts of social medicine, health management, primary  
118       health care, health statistics, and epidemiology et al.) and focus group (1 project leader, 1  
119       professor of epidemiology and health statistics, 1 health insurance administration, and 4  
120       graduate students), was used to construct the model. With reference to the literature, we  
121       refined the "perceived quality" variable to five first-order latent variables, namely, overall  
122       quality, information quality, service quality, policy quality, and institution quality. The draft  
123       model consisted of 11 latent variables, including public expectations (PE), perceived quality  
124       (PQ) (which was a second-order latent variable and consisted of the five first-order latent  
125       variables), perceived value (PV), public satisfaction (PS), public complaints (PC), and public  
126       trust (PT). Public expectation, perceived quality, and perceived value were the explanatory  
127       variables of public satisfaction, while public complaints and public trust were the outcome  
128       variables of public satisfaction. Consistent with Peng's research, we used the term "public" as  
129       an alternative of the term "customer" and changed "customer loyalty" to "public trust". Then,  
130       an initial draft variable pool (29 variables) was generated (see Table 1). Since the path of

131 public complaints on public trust was not significant according to Peng's research, this path  
132 was not set in the initial draft SIM-URRBMI.

133

134

135

136 **Table 1** The latent and measurement variables of the initial SIM\_URRBMI

Latent variable	First-order latent variable	Measurement variables
Public Expectation		Overall expectations (PE1) Expectations of URRBMI to ensure basic medical needs (PE2) Expectations of URRBMI to ensure personalization medical needs (PE3)
	Overall quality	Overall evaluation of quality (PQ1)
	Information quality	Availability of related information (PQ2) The convenience of obtaining relevant information (PQ3)
	Service quality	Service attitude of staff (PQ4) Clarity of staff's interpretation of relevant policies (PQ5) Degree to meet the basic health insurance needs (PQ6) Degree to meet the personalized health insurance needs (PQ7) Satisfaction with the payment level (PQ8)
Perceived Quality	Policy quality	Satisfaction with the scope of reimbursement (PQ9) Satisfaction with the reimbursement proportion (PQ10) Satisfaction with the deductible level (PQ11) Satisfaction with the capitation (PQ12) Service quality of community or village committees for enrollment (PQ13)
	Quality of institutions	Quality of payment bank or online payment platform (PQ14) Service quality of authorized medical institution (PQ15) Service quality of reimbursement in medical institution (PQ16) Service quality of authorized pharmacy (PQ17)
Perceived Value		Quality according to the payment (PV1) Payment according to the quality (PV2)
Public Satisfaction		Overall satisfaction (PS1) Service satisfaction to meets expectations (PS2)

	Distance close to the ideal medical insurance (PS3)
Public Complaint	Complained informally about URRBMI (PC1)
	Complained formally about URRBMI (PC2)
Public Trust	Willingness to positively evaluate (PT1)
	Willingness to recommend (PT2)

137

138       Based on expert consultation and literature review, the following path hypotheses were

139       established in the study (Fig 1):

140       **Fig1 The initial draft of SIM\_URRBMI**

141       Hypothesis 1: Public expectation has a direct positive impact on perceived quality,

142       perceived value, and public satisfaction;

143       Hypothesis 2: Perceived quality has a direct positive impact on perceived value and

144       public satisfaction;

145       Hypothesis 3: Perceived value has a direct positive impact on public satisfaction;

146       Hypothesis 4: Public satisfaction has a direct positive impact on public trust;

147       Hypothesis 5: Public satisfaction has a direct negative impact on public complaints.

148       **Phase 2**

149       Classic Test Theory (CTT) combined with Item Response Theory (IRT) was applied for

150       the measurement variable selection. If a measurement variable reached 3 or more deletion

151       criteria, it was considered for deletion. CTT is based on the true score model (true score plus

152       error), and its selection methods and criteria are shown in Table 2. The basic idea of IRT is to

153       use a mathematical function to characterize the relationship between the test response; in this

154       study,  $\theta$  refers to subject satisfaction with URRBMI. For the graded response model (GRM)

155       in this study, the discrimination parameter ( $a$ ), Threshold Parameter ( $b$ ), and Item Information

156 Function (*IIF*) were used to evaluate the quality of measurement variables. In addition, the  
157 Test Information Function (*TIF*) is a linear cumulative measurement of each item *IIF*; that is,  
158 the test information function on the  $\theta$  value is equal to the sum of its *IIFs*. It is generally  
159 considered that the quality of the model is good when the total of the information function is  
160 greater than 25, while the quality is considered poor when the information function is less  
161 than 16[15-16]. Since the initial draft of the SIM-URRBMI contains a total of 29  
162 measurement variables, the average information function amount of each measurement  
163 variable required ( $\bar{I}$ )  $\geq 0.55$  (16 / 29). The deletion criteria based on IRT are as follows: (1) *a*  
164 is less than 0.3; (2) *b* is out of range of (-4,4); and (3)  $\bar{I}$  is less than 0.55.

165

**Table 2 Item selection methods and criteria based on CTT**

Methods	Criteria	Main test aspects
Floor effect	If 20% or more individuals choose the lowest score value, then delete it.	Sensitivity
Ceiling effect	If 20% or more individuals choose the highest score value, then delete it.	Sensitivity
t-test	The measurement variables with no statistically significant difference were deleted ( $\alpha=0.05$ ) between 27% of the individual groups with the highest and lowest satisfaction scores.	Sensitivity or discrimination
Cronbach's $\alpha$ coefficient	If there is a large increase in the Cronbach's $\alpha$ of the dimension after the removal of an observed variable, then delete it.	Internal consistency
Correlation analysis 1	If the Pearson correlation coefficient of each measurement variable score and its corresponding dimension score is less than 0.6, then delete it.	Sensitivity or representation
Correlation analysis 2	If the Pearson correlation coefficient of each observed variable score and the initial model total score is less than 0.4, then delete it.	Representation
Stepwise regression	Using the total score of each latent variable as the dependent variable, the stepwise regression analysis is performed with candidate measurement variables as independent variables ( $sle = 0.05$ , $sls = 0.10$ ). The measurement variable excluded from the model will be deleted.	Importance
Factor analysis	Measurement variables with a factor load less than 0.4 will be deleted.	Independence

166

## 167 **Model validation**

168 The model validation includes two parts. The major reliability and validity tests were  
169 used to evaluate the measurement model. The value of Cronbach's  $\alpha$  coefficient and the  
170 composite reliability coefficient were greater than 0.7. It is considered to indicate good  
171 reliability, and higher values indicate greater homogeneity. Outer loading ( $\geq 0.708$ ) and  
172 Average Variance Extraction (AVE) ( $\geq 0.5$ ) were used to assess the convergence validity, while  
173 the Cross Loading and Fornell-Larcker criteria were used to assess discriminant validity.

174 Good discriminant validity means that the outer loading of a measurement variable on its  
175 corresponding latent variable is greater than its cross loadings on other latent variables, or the  
176 square root of the AVE of each latent variable is greater than its linear correlation coefficients  
177 with any other latent variables. Path analysis of the latent variables and model prediction were  
178 used to evaluate the structural model [11]. The relationship between latent variables was  
179 evaluated by estimating direct effect, indirect effect, and total effect. The model prediction  
180 assessment indicators were the adjusted coefficient of determination ( $R_{adj}^2$ ) (In general,  $R_{adj}^2$   
181 values of 0.25, 0.50, and 0.75 for target constructs are considered weak, medium, and  
182 substantial, respectively.), the effect size ( $f^2$ ) (Results of 0.02, 0.15, and 0.35 are interpreted  
183 as small, medium, and large  $f^2$  effect sizes, respectively.), and the predictive relevance ( $Q^2$ )  
184 (The path model has predictive relevance for a selected endogenous construct if the  $Q^2$  value  
185 is above zero).

186 **Statistical analysis**

187 Second-order Partial Least Squares Structural Equation Model (PLS-SEM) was used to  
188 test the model's reliability and validity, and PLS-SEM Multigroup Analysis (PLS-MGA) was  
189 used to explore the heterogeneity between rural and urban participants. A two-tailed  $P \leq 0.05$   
190 was considered statistically significant. To be consistent with the international customer  
191 satisfaction index, a one hundred percentage point system was used to express the public  
192 satisfaction index (PSI).

193 The descriptive statistics and CTT analyses were conducted using SPSS 18.0. IRT  
194 analysis was performed using Mplus version 7.0. PLS-SEM and PLS-MGA analyses were  
195 conducted with SmartPLS 3.0.

## 196      **Results**

### 197      **Samples**

198      A total of 1909 participants who had insurance for their children were included in the  
199      primary data analysis. Table 3 shows the sociodemographic information of pupils and main  
200      decision makers for the model construction group and model validation group separately. For  
201      pupils, the mean age was  $9.26 \pm 1.75$ , and the frequency of male (945, 49.89%) and female  
202      students (949, 50.11%) were similar. For the main decision makers, the mean age was  
203       $37.13 \pm 5.90$  years; most of decision makers were parents of pupils (1845, 96.95%), especially  
204      mothers (1196, 62.85%). Table 3 also indicates that there were no significant differences  
205      between the two groups for both pupils and main decision makers , suggesting that the two  
206      groups were homogenous. The measurement scores and comparative analysis between the  
207      two groups are shown in S1 Table .

208      **Table 3 Sociodemographic information of the pupils and the main decision makers for**  
209      **their BMI**

Variables	Model construction	Model validation	All	Statistics	P
	group ( n=574 )	group ( n=1335 )	( n=1909 )		
<b>Pupils</b>					
Age, Mean $\pm$ SD	9.26 $\pm$ 1.78	9.27 $\pm$ 1.74	9.26 $\pm$ 1.75	0.553	0.580
Sex, n (%)				2.610	0.106
male	299 (52.73%)	646 (48.68%)	945 (49.89%)		
female	268 (47.27%)	681 (51.32%)	949 (50.11%)		
Sum	567 (100.00%)	1327 (100.00%)	1894 (100.00%)		
Health condition <sup>a</sup> , n (%)				0.957	0.339
Very good	323 (56.77%)	728 (54.99%)	1051 (55.52%)		
Good	222 (39.01%)	529 (39.95%)	751 (39.67%)		
General	24 (4.22%)	67 (5.06%)	91 (4.81%)		
Sum	569 (100.00%)	1324 (100.00%)	1893 (100.00%)		
<b>Main decision maker</b>					

<b>Age, Mean ± SD</b>	37.24±5.79	37.07±5.88	37.13±5.90	0.121	0.904
<b>Relationship with pupils, n (%)</b>				1.968	0.579
Father	192 (33.62%)	457 (34.31%)	649 (34.10%)		
Mother	360 (63.05%)	836 (62.76%)	1196 (62.85%)		
Others	19 (3.33%)	39 (3.93%)	58 (3.05)		
Sum	571 (100.00%)	1332 (100.00%)	1903 (100.00%)		
<b>Sex, n (%)</b>				0.055	0.815
male	200 (35.03%)	474 (35.59%)	674 (35.42%)		
female	371 (64.97%)	858 (64.41%)	1229 (64.58%)		
Sum	571 (100.00%)	1332 (100.00%)	1903 (100.00%)		
<b>Marital status<sup>b</sup>, n (%)</b>				—	0.764
Married	552 (97.01%)	1275 (96.37%)	1827 (96.56%)		
Divorced	16 (2.81%)	40 (3.03%)	56 (2.96%)		
Never Married	0 (0.00%)	4 (0.30%)	4 (0.22%)		
Others	1 (0.18%)	4 (0.30%)	5 (0.26%)		
Sum	569 (100.00%)	1323 (100.00%)	1892 (100.00%)		
<b>Education<sup>a</sup>, n (%)</b>				3.230	0.520
Junior high school and below	163 (29.00%)	372 (28.38%)	535 (28.56%)		
High or technical secondary school	239 (42.53%)	583 (44.47%)	822 (43.89%)		
Junior college	94 (16.73%)	187 (14.26%)	281 (15.00%)		
Bachelor	57 (10.14%)	153 (11.67%)	210 (11.22%)		
Master or above	9 (1.60%)	16 (1.22%)	25 (1.33%)		
Sum	562 (100.00%)	1311 (100.00%)	1873 (100.00%)		

210           <sup>a</sup>Kruskal-Wallis *H* test   <sup>b</sup>Fisher's exact test

## 211   Model construction

212       S2 Table provided GRM estimations of the discrimination (*a*) and threshold parameters

213       (*b1-b9*) as well as their standard errors (SE). Table 4 shows the results of measurement

214       variable selection with the CTT and IRT methods. Using these methods, the ceiling effect

215       selected 5 variables (PQ1, PQ14, PE2, PT1 and PT2), Cronbach's *α* coefficient selected 1

216       variable (PS3), *a* selected 1 variable (PC1), *b* selected 9 variables (PQ9-PQ12, PQ16, PV1,

217       PV2, PS1 and PS2), and  $\bar{I}$  selected 4 variables (PQ1, PQ2, PC1, and PC2).

218

219 **Table 4 The results of measurement variable selection with the methods of CTT and IRT**

Variables	CTT						IRT			Selecte d	
	Floor effect	Ceiling effect	<i>t-test</i>	Correlation analysis 1	Correlation analysis 2	Cronbach's $\alpha$ coefficient	Stepwise regression	Factor analysis	<i>a</i>	<i>b</i>	
										$\bar{I}$	
PQ1		x							x		2
PQ2									x		1
PQ3											0
PQ4											0
PQ5											0
PQ6											0
PQ7											0
PQ8											0
PQ9									x		1
PQ10									x		1
PQ11									x		1
PQ12									x		1
PQ13											0
PQ14		x									1
PQ15											0
PQ16									x		1
PQ17											0
PE1											0
PE2		x									1
PE3											0
PV1									x		1
PV2									x		1
PS1									x		1
PS2									x		1

PS3		x		1
PC1	—	—	x x	2
PC2	—	—	x	1
PT1	x			1
PT2	x			1

220 x Suggested deletion

221 Overall, no item reached deletion criteria 3 or more times. However, it can be seen that  
222 the  $\bar{I}$  of PC2 is 0.005, suggesting that its contribution to the model is near 0, and its  
223 discrimination is also low ( $\alpha$  is 0.351). In addition, with China's social and cultural  
224 backgrounds, most enrollees will not formally complain even if they are dissatisfied. Thus,  
225 the PC2 was deleted, and the remaining 28 variables were used to construct the  
226 SIM-URRBMI (Fig 2).

227 **Fig2 The final version of SIM\_URRBMI**

228 **Model validation**

229 **Reliability**

230 The Cronbach's  $\alpha$  coefficient for each of the latent variables ranged from 0.772 to 1.000,  
231 and the composite reliability coefficients ranged from 0.897 to 1.000, both of which are  
232 greater than 0.7, providing evidence that the revised SIM-URRBMI is reliable (see Table 5).

233 **Table 5 The reliability of the SIM-UERRBMI**

Latent variables	Cronbach's $\alpha$ coefficient	Composite reliability coefficient
PQ	0.956	0.960
PE	0.918	0.948
PV	0.895	0.950
PS	0.867	0.919
PC	1.000	1.000
PT	0.772	0.897
PQ_overall	1.000	1.000
PQ_information	0.796	0.907

PQ_service	0.845	0.928
PQ_policy	0.933	0.946
PQ_institution	0.926	0.944

234 **Validity**

235 The results of the convergent and discriminant validity of the revised SIM-URRBMI are  
236 summarized in Tables 6 - 8. The *AVEs* of all latent variables were greater than 0.5, and the  
237 outer factor loadings of all the measurement variables ranged from 0.726 to 1.000, except  
238 that the second-order external factor loadings for the PQ1, PQ2, and PQ3 variables (0.692,  
239 0.595 and 0.684, respectively) was slightly below the critical value of 0.7, showing good  
240 convergent validity. All of the outer loadings of the measurement variables on their  
241 corresponding latent variable were greater than their cross loadings on the other latent  
242 variables. The square roots of the *AVEs* for each latent variable were greater than the linear  
243 correlation of the latent variable with the other structures, except for the perceived quality  
244 variable, which has a slightly higher correlation with perceived value and public  
245 satisfaction(Table 8). The results indicate good discriminant validity. All but one of the paths  
246 were statistically significant and positive; the effect of public satisfaction on the public  
247 complaints was negative, which is in line with the model theory framework. Table 9 and S3  
248 Table show the result of the direct effects and the collinearity assessment of the  
249 SIM-URRBMI. Furthermore, the greatest effect of the five first-order latent variables on  
250 perceived quality was medical insurance policy quality (0.472) followed by institution quality  
251 (0.354) (see Table 9). All variance inflation factor (*VIF*) values are clearly below the threshold  
252 of 5. Therefore, collinearity among the predictor constructs is not an issue in the structural  
253 model. The two paths with the greatest direct effect were perceived quality on perceived value  
254 (0.676) and public satisfaction on public trust (0.634). S4 Table indicates the indirect effects

255 of the revised SIM-URRBMI.

256

**Table 6 The convergent validity of the SIM-URRBMI**

Latent variables	Measurement variables	Outer factor loadings	Outer factor loadings (second order)	AVE
PQ	—	—	—	0.590
PQ_overall	PQ1	1.000	0.692	1.000
PQ_information	PQ2	0.898	0.595	0.830
	PQ3	0.924	0.684	
PQ_service	PQ4	0.930	0.756	0.866
	PQ5	0.931	0.761	
	PQ6	0.788	0.757	
PQ_policy	PQ7	0.783	0.731	0.716
	PQ8	0.796	0.750	
	PQ9	0.901	0.833	
PQ_policy	PQ10	0.903	0.836	
	PQ11	0.884	0.832	
	PQ12	0.859	0.801	
	PQ13	0.856	0.783	
	PQ14	0.836	0.726	
PQ_institution	PQ15	0.912	0.839	0.759
	PQ16	0.920	0.844	
	PQ17	0.869	0.794	
PE	PE1	0.921	—	
	PE2	0.943	—	0.859
	PE3	0.916	—	
PV	PV1	0.951	—	0.905
	PV2	0.952	—	
PS	PS1	0.925	—	
	PS2	0.928	—	0.791
	PS3	0.811	—	
PC	PC1	1.000	—	1.000
PT	PT1	0.887	—	0.813
	PT2	0.916	—	

257

**Table 7 Outer factor loadings and cross loadings of measurement variables for  
theSIM-URRBMI**

Variables	PQ_overal	PQ_information	PQ_service	PQ_policy	PQ_institution	PQ_(second order)	PE	PV	PS	PC	PT
	1										
PQ1	<b>1.000</b>	0.499	0.582	0.586	0.597	<b>0.692</b>	0.427	0.561	0.584	-0.224	0.474
PQ2	0.428	<b>0.898</b>	0.496	0.515	0.439	<b>0.595</b>	0.287	0.444	0.460	-0.118	0.305
PQ3	0.479	<b>0.924</b>	0.619	0.588	0.529	<b>0.684</b>	0.333	0.482	0.515	-0.132	0.375
PQ4	0.577	0.544	<b>0.930</b>	0.595	0.703	<b>0.756</b>	0.440	0.565	0.597	-0.216	0.467
PQ5	0.507	0.603	<b>0.931</b>	0.645	0.645	<b>0.761</b>	0.434	0.568	0.604	-0.204	0.416
PQ6	0.533	0.546	0.610	<b>0.788</b>	0.584	<b>0.757</b>	0.448	0.596	0.624	-0.196	0.439
PQ7	0.442	0.537	0.566	<b>0.783</b>	0.556	<b>0.731</b>	0.389	0.540	0.591	-0.190	0.392
PQ8	0.505	0.478	0.514	<b>0.796</b>	0.612	<b>0.750</b>	0.455	0.671	0.633	-0.267	0.524
PQ9	0.521	0.516	0.578	<b>0.901</b>	0.666	<b>0.833</b>	0.394	0.674	0.698	-0.260	0.476
PQ10	0.512	0.500	0.565	<b>0.903</b>	0.681	<b>0.836</b>	0.421	0.680	0.708	-0.238	0.520
PQ11	0.494	0.522	0.570	<b>0.884</b>	0.690	<b>0.832</b>	0.409	0.681	0.684	-0.262	0.504
PQ12	0.463	0.502	0.544	<b>0.859</b>	0.658	<b>0.801</b>	0.366	0.652	0.672	-0.280	0.481
PQ13	0.490	0.456	0.675	0.644	<b>0.856</b>	<b>0.783</b>	0.493	0.590	0.650	-0.213	0.517
PQ14	0.507	0.418	0.583	0.573	<b>0.836</b>	<b>0.726</b>	0.478	0.565	0.597	-0.193	0.503
PQ15	0.546	0.505	0.653	0.706	<b>0.912</b>	<b>0.839</b>	0.521	0.678	0.735	-0.237	0.568
PQ16	0.553	0.507	0.663	0.706	<b>0.920</b>	<b>0.844</b>	0.534	0.693	0.747	-0.255	0.575
PQ17	0.526	0.458	0.608	0.671	<b>0.869</b>	<b>0.794</b>	0.535	0.629	0.671	-0.207	0.539
PE1	0.397	0.332	0.449	0.448	0.559	0.534	<b>0.921</b>	0.537	0.546	-0.083	0.486
PE2	0.391	0.297	0.444	0.437	0.542	0.517	<b>0.943</b>	0.547	0.541	-0.074	0.491
PE3	0.400	0.322	0.411	0.468	0.520	0.524	<b>0.916</b>	0.544	0.531	-0.052	0.454
PV1	0.540	0.484	<b>0.577</b>	0.726	0.685	0.755	0.556	<b>0.951</b>	0.795	-0.293	0.601
PV2	0.528	0.485	0.581	0.721	0.684	0.752	0.557	<b>0.952</b>	0.807	-0.264	0.589
PS1	0.581	0.493	0.629	0.741	0.754	0.797	0.559	0.831	<b>0.925</b>	-0.297	0.598
PS2	0.544	0.513	0.616	0.743	0.729	0.787	0.528	0.791	<b>0.928</b>	-0.276	0.563
PS3	0.419	0.420	0.462	0.583	0.571	0.615	0.461	0.604	<b>0.811</b>	-0.212	0.526
PC1	-0.275	-0.169	-0.277	-0.351	-0.310	-0.351	-0.093	-0.359	-0.364	<b>1.000</b>	-0.251
PT1	0.389	0.298	0.390	0.472	0.520	0.520	0.456	0.509	0.530	-0.142	<b>0.887</b>
PT2	0.462	0.376	0.461	0.542	0.587	0.600	0.472	0.613	0.607	-0.223	<b>0.916</b>

259 The bold number in the table is the external factor loadings of the measurement variable on its

260 corresponding latent variable, and the rest are the cross loadings.

261

**Table 8 The Fornell-Larcker standard for the SIM-URRBMI**

	PQ	PQ_overall 1	PQ_information	PQ_service	PQ_policy	PQ_institution	PE	PV	PS	PC	PT
PQ	<b>0.768</b>										
PQ_overall	0.692	<b>1.000</b>									
PQ_Information	0.705	0.499	<b>0.911</b>								
PQ_service	0.815	0.582	0.616	<b>0.931</b>							
PQ_policy	0.937	0.586	0.608	0.666	<b>0.846</b>						
PQ_institution	0.908	0.597	0.534	0.724	0.753	<b>0.879</b>					
PE	0.567	0.427	0.342	0.470	0.486	0.583	<b>0.927</b>				
PV	0.792	0.561	0.510	0.609	0.760	0.720	0.585	<b>0.951</b>			
PS	0.830	0.584	0.536	0.646	0.780	0.776	0.582	0.842	<b>0.890</b>		
PC	-0.286	-0.224	-0.138	-0.226	-0.286	-0.252	-0.076	-0.293	-0.297	<b>1.000</b>	
PT	0.623	0.474	0.376	0.475	0.565	0.616	0.515	0.625	0.633	-0.205	<b>0.902</b>

262 The bolded numbers in the diagonal of the table are the square roots of AVE for each latent

263 variable

264

**Table 9 The direct effects in the SIM-URRBMI**

Direct path	Coefficient	SE	t	P	95%CI	
					Lower	Upper
PE→PQ	0.568	—	—	—	—	—
PE→PV	0.204	0.027	7.372	<0.001	0.145	0.253
PE→PS	0.070	0.021	3.362	0.001	0.034	0.119
PQ→PV	0.676	0.021	31.964	<0.001	0.639	0.720
PQ→PS	0.421	0.028	14.960	<0.001	0.361	0.474
PV→PS	0.467	0.032	14.698	<0.001	0.404	0.528
PS→PC	-0.243	0.020	11.847	<0.001	-0.280	-0.202
PS→PT	0.634	0.020	31.305	<0.001	0.596	0.674
PE→PQ_overall	0.428	0.027	15.651	<0.001	0.373	0.477
PE→PQ_information	0.344	0.026	12.981	<0.001	0.290	0.393
PE→PQ_service	0.471	0.026	18.247	<0.001	0.413	0.518
PE→PQ_policy	0.487	0.025	19.415	<0.001	0.437	0.534
PE→PQ_institution	0.584	0.023	25.431	<0.001	0.534	0.624
PQ_overall→PQ	0.069	0.002	36.692	<0.001	0.066	0.073
PQ_information→PQ	0.110	0.003	31.786	<0.001	0.103	0.117
PQ_service→PQ	0.135	0.003	47.898	<0.001	0.130	0.142
PQ_policy→PQ	0.472	0.006	76.267	<0.001	0.462	0.485
PQ_institution→PQ	0.354	0.005	64.894	<0.001	0.343	0.364

265 The total effect of the 14 paths of the SIM-URRBMI were statistically significant,  
266 showing that the model was constructed properly (see Table 10). The three strongest paths  
267 were (1) perceived quality on public satisfaction (with a total effect of 0.737, a direct effect of  
268 0.421 and an indirect effect of 0.316); (2) perceived quality on perceived value (with a total  
269 effect of 0.676);and (3) public satisfaction on public trust (with a total effect of 0.634). The  
270 largest total effect of the five explanatory variables of public trust was calculated for public  
271 satisfaction (0.634), followed by perceived quality (0.467). Among the three explanatory  
272 variables of public satisfaction, the greatest total effect was from perceived quality (0.737).

273 **Table 10 The results of the total effect analysis for the SIM-URRBMI**

Total path	Coefficient	SE	t	P	95%CI	
					Lower	Upper
PE→PQ <sup>a</sup>	0.568	0.022	25.377	<0.001	0.519	0.609
PE→PQ	0.588	0.025	23.122	<0.001	0.535	0.637
PE→PS	0.584	0.024	23.953	<0.001	0.535	0.630
PE→PC <sup>b</sup>	-0.142	0.013	10.963	<0.001	-0.168	-0.116
PE→PT <sup>b</sup>	0.371	0.023	16.142	<0.001	0.322	0.409
PQ→PV <sup>a</sup>	0.676	0.021	31.964	<0.001	0.639	0.720
PQ→PS	0.737	0.018	41.270	<0.001	0.699	0.771
PQ→PC <sup>b</sup>	-0.179	0.016	10.870	<0.001	-0.210	-0.143
PQ→PT <sup>b</sup>	0.467	0.018	25.732	<0.001	0.430	0.502
PV→PS <sup>a</sup>	0.467	0.032	14.698	<0.001	0.404	0.528
PV→PC <sup>b</sup>	-0.113	0.013	9.010	<0.001	-0.139	-0.090
PV→PT <sup>b</sup>	0.297	0.023	13.161	<0.001	0.252	0.336
PS→PC <sup>a</sup>	-0.243	0.020	11.847	<0.001	-0.280	-0.202
PS→PT	0.634	0.020	31.305	<0.001	0.596	0.674

274 <sup>a</sup>indicates that only the direct path was set, <sup>b</sup>indicates that only the indirect path was set

275

276 The adjusted coefficients of determination ( $R_{adj}^2$ ) of each endogenous latent variable  
277 ranged from 0.088 to 1.000. The  $R_{adj}^2$  values were 0.783 and 0.400 for public satisfaction  
278 and public trust, respectively, indicating a strong and moderate prediction. The effect size ( $f^2$ )

279 ranged from 0.014 to 0.903; the  $f^2$  values for perceived quality on predicted perceived value,  
280 the perceived value on predicted public satisfaction, and public satisfaction on predicted  
281 public trust were 0.903, 0.351 and 0.667, respectively; all were greater than 0.35 indicating a  
282 strong predictive effect. The predictive relevance ( $Q^2$ ) of all endogenous latent variables was  
283 greater than 0, ranging from 0.105 to 0.585, showing a certain predictive relevance for all five  
284 endogenous latent variables (see Table 11).

285 **Table 11 The results of predictive ability evaluation for the SIM-URRBMI**

	<b>PE</b>	<b>PQ</b>	<b>PV</b>	<b>PS</b>	<b>PC</b>	<b>PT</b>
$R^2_{adj}$	—	1.000	0.654	0.783	0.088	0.400
The predictive correlation ( $Q^2$ )	—	0.550	0.567	0.585	0.105	0.311
PE	—	—	0.080	0.014	—	—
PQ	—	—	0.903	0.290	—	—
The effect scale ( $f^2$ )	PV	—	—	0.351	—	—
PS	—	—	—	—	0.097	0.667
PC	—	—	—	—	—	—
PT	—	—	—	—	—	—

286 **Model application**

287 The SIM-URRBMI satisfaction index scores were 57.29 and 58.17 for rural and urban  
288 participants, respectively. PLS-MGA showed that the total effect of public expectation on  
289 public satisfaction and public complaints was higher for citizens than for rural residents  
290 (0.627 vs 0.534,  $P=0.037$ ; -0.166 vs -0.114,  $P=0.042$ ). The total effect of perceived quality on  
291 public trust for citizens was lower than that of rural residents (0.429 vs 0.514,  $P=0.023$ ) (see  
292 Table 12).

293

294

295

296 **Table12** Comparison of total effects between urban and rural residents' among latent variables

<b>Total path</b>	<b>Rural areas</b>		<b>Urban areas</b>		<b>Rural vs Urban</b>		
	$p^{(1)}$	$se(p^{(1)})$	$p^{(2)}$	$se(p^{(2)})$	$ p^{(1)} - p^{(2)} $	$t$	$P$
PE→PQ	0.548	0.034	0.588	0.028	0.042	0.954	0.340
PE→PQ	0.568	0.036	0.610	0.032	0.040	0.834	0.404
PE→PS	0.534	0.036	0.627	0.029	0.095	2.087	0.037
PE→PC	-0.114	0.016	-0.166	0.020	0.052	2.034	0.042
PE→PT	0.360	0.033	0.378	0.029	0.019	0.442	0.659
PQ→PV	0.697	0.032	0.662	0.030	0.032	0.724	0.469
PQ→PS	0.762	0.025	0.714	0.025	0.048	1.380	0.168
PQ→PC	-0.164	0.024	-0.189	0.023	0.026	0.773	0.440
PQ→PT	0.514	0.027	0.429	0.025	0.083	2.277	0.023
PV→PS	0.499	0.042	0.448	0.043	0.056	0.920	0.358
PV→PC	-0.107	0.018	-0.119	0.017	0.011	0.442	0.659
PV→PT	0.336	0.032	0.270	0.030	0.069	1.589	0.112
PS→PC	-0.214	0.029	-0.265	0.029	0.051	1.229	0.219
PS→PT	0.674	0.028	0.602	0.029	0.071	1.741	0.082

297

## 298 Discussion

299 China's medical reform has entered a critical period, there is an urgent need to understand  
300 participants' satisfaction with BMI and to determine the main aspects that need improvement.  
301 This work will provide important evidence for the performance improvement of the insurance  
302 schemes and provide a critical reference for reform [4]. After more than 60 years of reform  
303 and construction, China has finally established a basic medical insurance system that covers  
304 the entire population for basic medicine. To meet participants' growing basic need for a better  
305 life, it is necessary to continuously improve the quality of universal medical insurance from  
306 the perspective of the insured [17].

307 Moreover, the quality of medical insurance determines the satisfaction of the insured to a  
308 large extent. The original SIM-URRBMI developed by Peng et al has provide a standard  
309 measurement of satisfaction, but could not provide detailed information for the diagnosis of  
310 poor quality as there were few detailed quality items. This study constructed the perceived  
311 quality variable as a second-order latent variable based on the literature review and expert  
312 consultation. The second-order SEM can not only reduce the number of structural model  
313 relationships but also make the model path more concise and easier to understand. Most  
314 importantly, it can measure the quality perception of URRBMI participants more accurately.  
315 This characteristic will greatly help identify the aspects of URRBMI that are in urgent need of  
316 improvement [18]. The results of the predictive ability evaluation showed that the  $R_{adj}^2$  of  
317 PS was 0.783, which is larger than the  $R^2$  of PS in the original version of SIM-URRBMI  
318 constructed by Peng et al. (0.63). From this point of view, it appears that the construction of  
319 PQ as a second-order latent variable in this study increased the predictive power of public

320 satisfaction. This also verifies our hypothesis that the quality-refined model can measure PS  
321 more accurately.

322 Additionally, this study also tested the path effect differences between rural and urban  
323 participants; thus, it can not only reveal population heterogeneity in structural models but also  
324 help identify the preferred path for quality and satisfaction improvement in a certain subgroup  
325 population.

326 The measurement variables selection is a key step in model construction. CTT and IRT  
327 complementary each other, combining both methods in measurement variable selection will  
328 provide more stable and reliable choices and benefits for model construction. The result of  
329 measurement variable selection revealed that the measurement variable PC2 (formal  
330 complaint) had poor performance. It is consistent with Chinese cultural backgrounds that the  
331 public seldom choose this method to express their dissatisfaction with public services.  
332 Deletion of PC2 suggested that this study had a good rationale for its measurement variable  
333 selection methods [19].

334 The Cronbach's  $\alpha$  coefficients and composite reliability coefficients for all latent  
335 variables were greater than 0.7, suggesting good internal consistency. except for the  
336 perceived quality variable, which was slightly less than the linear correlation coefficient  
337 between the perceived value and public satisfaction variables (One possible reason is that the  
338 second-order external factor loadings for the PQ1-PQ3 variables were less than 0.7), all the  
339 other results indicate that the model has good convergent validity and discriminant  
340 validity. The total effect of perceived quality on public satisfaction is 0.737, which has the  
341 largest total effect among all latent variables, indicating that PQ is the most important factor

342 affecting public satisfaction, which is consistent with the research conducted by Lang Xu and  
343 Fornell C in the medical insurance field and Johnson MD in other research fields [20-21].  
344 First, from the results of the path coefficients between perceived quality and its first-order  
345 latent variables, it was seen that the quality of medical insurance policy had the greatest  
346 impact on PQ with a total effect of 0.472, suggesting that medical insurance policy had the  
347 greatest impact on public satisfaction. Second, PQ9 (basic coverage scope), PQ10  
348 (coinsurance), PQ11 (deductible) and PQ12 (capitation) had larger second-order factor  
349 loadings on PQ. Moreover, PQ9-PQ12 had larger discrimination parameters (2.4-2.8) from  
350 the IRT *IIF* results. According to Juxiang Qin's point of view, the discrimination parameter  
351 reflects the sensitivity, which can distinguish the sensitivity of all the participants' responses  
352 to the question (that is, whether the project is a concern shared by everyone), and the *IIF*  
353 result reflects that the insured residents are more concerned about the policies related to  
354 reimbursement [22]. Therefore, reimbursement policy for URRBMI has the greatest impact  
355 on the satisfaction of the insured residents, suggesting that the improvement of quality,  
356 especially the medical insurance reimbursement policy quality, is crucial to improving  
357 satisfaction with URRBMI.

358 The SIM-URRBMI satisfaction index score of the main family decision makers for the  
359 pupils' BMI in Changsha was 57.59, which was lower than the satisfaction scores of the  
360 insured farmers on the NCMS in 2014 (66.33) and the US federal government satisfaction  
361 index scores (69.70) in 2017 [21, 23]. The satisfaction index score can also be used as a  
362 vertical comparison tool for the satisfaction research for URRBMI assessment. The  
363 satisfaction index scores were 57.29 and 58.17 in the rural and urban areas, respectively,

364 indicating that the urban residents had a slightly higher satisfaction level of URRBMI in  
365 Changsha, which is consistent with the research conducted by Fan Yan et al.[24]. The  
366 PLS-MGA result showed that, compared to the rural residents, urban participants had a higher  
367 total effect of PE on both PS and PC, while having a lower total effect of PQ on PT. PE is a  
368 kind of expectation for future medical insurance policies based on past consumption  
369 experiences. The urban residents had better policy welfare than the rural residents before the  
370 consolidation. Thus, urban residents had higher expectations for URRBMI, and whether the  
371 expectation is close or far from the actual situation, they may have, accordingly, had higher  
372 positive effects (satisfaction) or negative effects (complaints) than rural residents.

373 There are also some limitations of this research. First, the sample population was the  
374 main family decision makers for pupils' URRBMI and did not involve the other URRBMI  
375 participants in the family (such as college students). The conclusion is limited to the entire  
376 insured population. It is hoped that different insured populations can be included in future  
377 stages to expand the applicability of the model. Second, this study is a cross-sectional study  
378 that did not collect longitudinal data and, thus, could not estimate trends in URRBMI  
379 satisfaction in the sample area.

## 380 **Conclusions**

381       The final version of the SIM-URRBMI consists of 28 measurement variables and 11  
382       latent variables. It is a reliable and valid satisfaction measurement tool with good prediction  
383       ability for public satisfaction and public trust. In addition, the model provides accurate  
384       assessment for perceived quality of URRBMI, which will greatly help for performance  
385       improvement.

## 386 **Acknowledgements**

387       The authors would like to thank all the participants, teachers, and leaders in the selected  
388       8 primary schools, Liuyang City Center For Disease Control and Prevention and Tianxin  
389       District Education Bureau.

## 390 **References:**

- 391       1. Jing LM, Chen R, Jing L, Qiao Y. Development and enrolee satisfaction with basic  
392       medical insurance in China: A systematic review and stratified cluster sampling survey.  
393       Int J Health Plann Manage. 2017;32(3):285-298.
- 394       2. Pan XF, Xu J, Meng Q. Integrating social health insurance systems in China. Lancet  
395       2016;387(10025):1274-1275.
- 396       3. Zhu K, Zhang L, Yuan S, Zhang X, Zhang Z. Health financing and integration of urban  
397       and rural residents' basic medical insurance systems in China. Int J Equity Health.  
398       2017;16(1):194.
- 399       4. Mohammed S, Sambo MN, Dong HJ. Understanding client satisfaction with a health  
400       insurance scheme in Nigeria: factors and enrollees experiences. Health Res Policy Syst.  
401       2011;9(1):1-8.

- 402 5. Duan JM. [Research on Satisfaction of Medical Insurance in Rural Immigrant  
403 Community in Guazhou County]. Gansu Agricultural University; 2016.Chinese.
- 404 6. Fornell C, Johnson M, Anderson EW, Cha J, Bryant BE. The American Customer  
405 Satisfaction Index: Nature, Purpose, and Findings. *J Mark.* 1996;60(4):7-18.
- 406 7. Wang Q, Mo XK, Peng LA. [Evaluation of a Satisfaction Index Model for Pupils with  
407 Basic Medical Insurance for rural and Urban residents]. *Chinese Journal of Health  
408 Statistics* 2015;32(1):37-40. Chinese.
- 409 8. Peng LA, Mo XK, Shi JC, Lei SY. [Construction of a Satisfaction Index Model for Pupils  
410 with Basic Medical Insurance for rural and Urban residents: Model Item Selection].  
411 *Chinese Journal of Health Statistics* 2013;30(1):532-535+539. Chinese.
- 412 9. Fornell C National Customer Satisfaction Barometer: The Swedish Experience. *J Mark.*  
413 1992;56(1):6-21.
- 414 10. Choi KS, Cho WH, Lee S, Lee H, Kim C. The relationships among quality, value,  
415 satisfaction and behavioral intention in health care provider choice: A South Korean  
416 study. *J Bus Res.* 2004;57(8):913-921.
- 417 11. Hair JF, Hult GTM, Ringle CM, Sarstedt M. *A Primer on Partial Least Squares Structural  
418 Equation Modeling (PLS-SEM)*. 2nd Ed ed: Thousand Oaks: SAGE; 2016.
- 419 12. Embretson B S E, Reise S P. *Item response theory for psychologists*. Mahwah: Lawrence  
420 Erlbaum, 2000.
- 421 13. Westland JC. Lower Bounds on Sample Size in Structural Equation Modeling. *Electron  
422 Commer Res Appl.* 2012;9(6):476-487.
- 423 14. Vinzi VE, Chin WW, Henseler J, Wang H. *Handbook of Partial Least Squares: Concepts,*

- 424        Methods and Applications. Berlin, Heidelberg, Springer; 2010.
- 425        15. Dani VV. Measuring Customer Satisfaction for F&B Chains in Pune Using ACSI Model.
- 426        Procedia Soc Behav Sci. 2014;133:465-472.
- 427        16. Qi SQ, Dai HQ. [Item Response Theory and Its Application Research]. Nanjing: Press of
- 428        Jiangxi education; 1992. Chinese.
- 429        17. Wang ZF. [Functional Positioning and Governance Mechanism of Medical
- 430        Insurance—Thoughts on Establishing National Medical Security Bureau]. Chin Heal
- 431        Insur. 2018(04):13-17. Chinese.
- 432        18. Ketchen, David J. A Primer on partial least squares structural equation modeling
- 433        (pls-sem). Long Range Plann. 2014;46(s1-2):184-185.
- 434        19. Sun ZQ, Xu YY. [Medical Statistics]. Fourth Edition ed. Beijing: People's Medical
- 435        Publishing House; 2014. Chinese.
- 436        20. Jonhson M, Gustafsson A, Andreassen TW, Lervik L, Cha J. The evolution and future of
- 437        national customer satisfaction index models. J Econ Psychol. 2001;22(2):217-245.
- 438        21. Xu L, Wang SL, Cao L. [Research on Satisfaction of new rural cooperative medical
- 439        system of farmers in Jiaozuo City of Henan Province]. Journal of Zhejiang Agricultural
- 440        Sciences. 2014(10):1643-1648. Chinese.
- 441        22. Qin JX. [Studies of Happiness Based on Item Response Theory]. Southwest University;
- 442        2016. Chinese.
- 443        23. ACSI 2017. ACSI Federal Government Report 2017 [Internet]. USA; 2017[cited 2
- 444        019 Mar 20]. Available from: <https://www.theacsi.org/news-and-resources/customer-satisfaction-reports/report-archive/acsi-federal-government-report-2017>.

446 24. Yan F, Ma GD, Pan JJ, Xu N, Gu Y, Sun YF. [Investigation on satisfaction of overall  
447 urban and rural residents after participating basic medical insurance]. St Sci Health.  
448 2018;32(01):39-41. Chinese.

449 **Supporting information**

450 **S1 Table. The measurement scores and comparative analysis between two groups of the**  
451 **initial draft of SIM\_URRBMI.**

452 **S2 Table. Item parameter estimation from the GRM.**

453 **S3 Table. The result of the collinearity assessment in the revised SIM\_URRBMI.**

454 **S4 Table. The indirect effects in the revised SIM\_URRBMI.**

455 **Funding**

456 This work was supported by the project of "13th five year plan" of Hunan Education  
457 Science in 2019 (XJK19BGD003), the China Community Health Services & Health  
458 Education Program (2014CC03), and education and teaching reform research project of  
459 Central South University in 2019 (2019jy146).



