**Tertiary hospitals or community clinics? An enquiry into the factors affecting patients’ choice for healthcare facilities in urban China**

**Abstract**

Most Chinese patients tend to seek primary care services in tertiary hospitals rather than community clinics, which has caused the malfunction of China’s multi-tiered health system. This study aimed to identify the factors that affected patients’ decisions when selecting healthcare facilities and use these to design policies to redirect patient flow from tertiary hospitals back to community clinics. A cross-sectional survey was conducted in Shanghai in 2017 using the best-worst scaling experiment. A total of 202 respondents were included in the analysis, including 97 community residents, 56 tertiary hospital patients, and 49 community clinic patients. The latter two had made their choices so the three samples were analyzed separately. Seven attributes were used in the experiment with each varying across two or three levels to accommodate the features of primary care services in Shanghai. The values of attribute levels were estimated using the mixed logit model. Relative importance across attributes and attribute levels were derived and policy simulations were undertaken to examine the impact of changing attribute levels on choices. Results suggest that the three samples are composed of different types of people who presented different preference patterns. The community residents are representative of the population and to them sufficient test and examinations, doctors with expertise, and good service attitude are the most desirable features for a healthcare facility while unfriendly practitioners and lengthy visit the least attractive features. In terms of attributes, friendliness of doctors and the availability of tests and examinations are the two major drivers in selecting healthcare facilities. Our findings provide supportive evidence and useful inputs to the ongoing healthcare reform in China.

**Key words:** Best-worst scaling; China healthcare reform; Community clinics; Patient choice; Primary care; Tertiary hospital

**JEL Classification:** C35, I18, P35

# **Introduction**

# Healthcare systems all over the world aim to provide high-quality healthcare services in an accessible, affordable, efficient, and sustainable way. An ideal pathway for achieving this aim is to provide healthcare services with a multi-tiered system, including primary care, secondary care, and tertiary care. Primary care is usually the first point of contact for people in need of healthcare, where the general practitioner (GP) makes referral to secondary or tertiary care when needed.

# In China, the government is establishing a multi-tiered medical system with a two-way referral mechanism, in order to allocate healthcare resources more efficiently by aligning them with clinical needs (State Council of China, 2009). The multi-tiered system intends to provide primary care in community clinics or health centers[[1]](#footnote-1), secondary care in district hospitals (known as secondary hospitals), and tertiary care in tertiary hospitals. In addition, the system generates referrals not only from primary care providers to hospitals, but also referrals in the reverse direction. Reverse referrals back to primary care enable patients to obtain some care services, such as rehabilitation and chronic disease management, at an affordable cost.

As there is no formal gatekeeper role, patients in China can access the system at any level and choose a healthcare facility freely. Hence, the success of this three-tiered hospital system crucially depends on whether patients choose a facility based on their healthcare needs. Unfortunately, Chinese patients often prefer accessing higher level hospitals directly, in particular tertiary hospitals, even for simple care needs (Wu & Lam, 2016). As a result, tertiary hospitals are often crowded and over utilized, while community clinics experience noticeable idle time (Zhao, Chen, & Liu, 2009).

In response to this problem, promoting primary care has become a major target of the latest healthcare reform launched in 2009. All levels of government in China have poured resources into primary care infrastructure construction, personnel training and other supporting programs (Yip, Hsiao, Chen et al., 2012). However, the investment has not enticed patients with primary care needs away from tertiary hospitals. In fact, the situation has worsened (Wu & Lam, 2016).

This phenomenon has generated substantial interest in understanding the reasons that have caused the failure and identifying potential solutions. Wu and Lam (2016) identified five interrelated causes including severe maldistribution of quality human resources, low educational attainment of primary care doctors, the public’s lack of trust in the quality of primary care, lack of a gatekeeping function, and the detrimental elements of health reform policies. A recent systematic review by Liu, Kong, Yuan et al. (2018) identified 45 studies that investigate the factors that affect patients’ choice of healthcare providers in China. These included: individual, provider, context[[2]](#footnote-2) and composite factors (combined individual, provider, and/or context attributes) and suggested that the provider factor ‘drug variety’ and ‘equipment’, and composite factor ‘perceived quality’, may push patients from lower levels towards higher levels.

These studies, however, are mostly content with examining the association between a factor and patient choice. Few studies have estimated the value attached to these factors and thus their relative importance, which is essential to understanding those factors that drive patients’ choice and facilitate more effective policy design. The only exception is Liu, Kong, and de Bekker-Grob (2019), which undertook a discrete choice experiment (DCE) to elicit the preference weights (or value) for several selected attributes/factors when choosing a healthcare facility in rural areas of Chongqing. They found improvements to drug availability, medical professional skill and equipment in the rural primary care system may redirect patient flow from higher level hospitals to primary level.

This study uses a best-worst scaling (BWS) experiment to assess patients’ choice for healthcare facilities in Shanghai when their care need is simple (mild disease). Shanghai was selected as an example of an urban area in China for two reasons. First, it is one of the biggest cities with its residents coming from all over the country. Second, its patients use high level healthcare facilities (in particular tertiary hospitals) much more than other parts of China, mainly because the city has many large hospitals with world-class technology and experts in almost any discipline (Yang & Yang, 2009). Indeed, a survey by Yu, Liu, and Huang (2007) found that almost half of the Shanghai respondents stated they would never choose community clinics under any circumstances.

Our paper makes several contributions. It elicits public preferences for healthcare facilities in urban China for the first time. BWS is a relatively new stated preference methodology, which does not only elicit relative importance across attributes but also across all attribute levels (Terry N Flynn, Louviere, Peters et al., 2007b). This allows us to obtain a full ranking of attribute levels, which can be of great value for policy design. We notably undertook a policy simulation task to examine potential channels on how to redirect patient flow from tertiary hospitals to community clinics for their primary care needs.

Our paper also adds to the broad literature on patient choice for healthcare providers from a unique perspective. This literature is focused on empirically testing the assumption that patient choice is influenced by the quality of healthcare providers and flexible choice may improve care quality by inducing competition between healthcare providers (Brekke, Gravelle, Siciliani et al., 2014). Such competition clearly exists among different levels of hospitals in China and is primarily the reason behind the underuse of primary care facilities (Wu & Lam, 2016). Since closing these facilities is not an option and nudges are limited, the most pragmatic way to change this situation might be to make these lower tier facilities more competitive, and this requires them to better meet patient preference which will be explored in this paper.

1. **Methods**

## *2.1 The BWS experiment*

Stated preference methods have become increasingly popular in economics research. One example is DCE, which is particularly useful for evaluating markets with restricted choice, where characteristics driving choices are unobserved, or where product attributes are not measured or lack variation (de Bekker‐Grob, Ryan, & Gerard, 2012). The BWS, devised by Finn and Louviere (1992) and introduced to healthcare research by McIntosh and Louviere (2002), is an alternative choice-based stated preference elicitation method. There are three types of BWS techniques (Louviere, Flynn, & Marley, 2015) and the second type, the profile case, has recently gained popularity in health economics research (T. N. Flynn, Louviere, Peters et al., 2007a). Our study used this profile case and we call it BWS hereafter.

BWS presents one profile at a time, constituted by several attributes of a healthcare service or product, with each taking a specific level. In a BWS choice task, respondents are asked to compare the attribute levels directly and select the best and worst ones when evaluating the healthcare service or product described in the profile. Through repeated choices, the values that respondents place on different attribute levels are elicited. BWS has two significant advantages over DCE. First, it compares attribute levels directly since they become the choice set, so respondents directly assess the relative attractiveness of levels across attributes. The task allows the relative utility of the levels of all attributes to be compared on the same scale, with a single level of a single attribute omitted as the base case. Second, it is less cognitively demanding than DCE, because respondents see only one profile at a time[[3]](#footnote-3) (Cheung, Wijnen, Hollin et al., 2016; Coast, Flynn, Natarajan et al., 2008; Terry N Flynn et al., 2007b).

## *2.2 Survey development*

A BWS experiment starts by defining the choice context. As this research is interested in patients’ preference for health care facilities when their care need is simple, we asked the respondents to imagine that they have one mild disease, e.g., common cold, influenza, fever, diarrhea, etc.

The second step and major focus of the BWS experiment is to identify the characteristics or attributes of the health care facility. A comprehensive literature review was first undertaken (up to March 2017) which led to four groups of attributes measuring cost, quality, access to services, and friendliness of doctors. Quality is measured by the type of physician, tests/examination availability and medication variety, while access to services is measured by travel time and total visit time. Literature review details are documented in the Appendix A.

This was followed by two semi-structured interviews (five participants each) in April 2017 and two focus group meetings (five participants each) in May 2017 to validate the attributes and finalize their levels. One interview and one focus group were conducted among patients at a tertiary hospital, while the other interview and focus group meeting were conducted among patients at a community clinic (details documented in the Appendix B). This process did not lead to additional attributes but helped us refine the definition and identify appropriate levels for the seven attributes (Table 1).

**Table 1. Attributes, levels and explanations**

|  |  |  |  |
| --- | --- | --- | --- |
| Attribute |  | Level | Explanation |
| Cost | 1 | CNY 100 | The out-of-pocket cost for medical and drug fees |
|  | 2 | CNY 325 |
|  | 3 | CNY 500 |
|  |  |  |  |
| Type of physician | 1 | General practitioner | Physicians are categorized as general practitioner and specialist to reflect the multi-tiered healthcare system |
|  | 2 | Specialist |
|  |  |  |  |
| Tests/examinations availability | 1 | Few (reference) | It reflects the general condition of equipment in a healthcare facility and therefore the capability of diagnosis |
|  | 2 | Many |
|  |  |  |  |
| Variety of medications | 1 | Limited | The availability of commonly used medicine in a healthcare facility |
|  | 2 | Great |
|  |  |  |  |
| Travel time to healthcare facility | 1 | 20 mins | The total time needed for travelling from home to the healthcare facility using public transportation |
|  | 2 | 40 mins |
|  | 3 | 60 mins |
|  |  |  |  |
| Total visit time in healthcare facility | 1 | 1 hour | The total time taken from entering a healthcare facility to leaving the facility. |
|  | 2 | 3 hours |
|  | 3 | 5 hours |
|  |  |  |  |
| Friendliness of doctors | 1 | Cold or rude | It reflects the service attitude of the doctors in a healthcare facility. |
|  | 2 | Friendly or considerate |

The cost attribute refers to the out-of-pocket cost (OOP) for medical and drug fees. In the interviews and focus group meetings, we asked participants to write down the OOP costs of their past ten physician visits for mild conditions. These numbers resulted in an empirical distribution of the OOP costs, which helped us draw three representative levels for the experiment. Likewise, the levels for physician type, tests/examination availability, medication variety, and staff friendliness were all specified to reflect the current healthcare services in Shanghai.

The travel time attribute was defined as the time needed to travel from home to a health care facility (one-way travel) using public transportation. Various similar measures, such as distance, were considered and extensively discussed during the interviews and focus groups, but travel time was deemed as the most appropriate attribute with little heterogeneity among respondents in their understanding. It is also widely used in the patient choice literature (Brekke et al., 2014).

Total visit time was used instead of waiting time, which is a more commonly considered attribute of patient choice (Brekke et al., 2014). In Chinese tertiary hospitals, patients are usually asked to take tests/examinations to assist physicians’ diagnosis. They have to visit different hospital departments and spend a substantial amount of time getting many tests done before finally seeing the doctor. This is also one main problem the healthcare reforms aim to resolve. Waiting time (the time spent waiting outside the doctor’s office) is relatively a smaller component of the total visit time and Chinese tertiary hospitals have already implemented interventions to minimize the length of waiting time (Sun, Lin, Zhao et al., 2017). Patients spend little talking to their doctors, typically only a few minutes. This is negligible compared to the other two components of the total visit time and it does not seem to be a pressing issue for the participants of our focus groups. It was the time spent visiting many different departments for tests and examinations that they mostly complained about. We therefore believe that the respondents would perceive this attribute as the time needed for seeing the doctor, a slightly broader definition of waiting time.[[4]](#footnote-4)

The tier of healthcare facility (e.g., tertiary hospital, community clinic) was not considered as an attribute as we intended to use different combinations of selected attribute levels to represent different tiers of healthcare facilities. In this case, including the tier attribute would induce correlations with the other attributes, which is not ideal for the experimental design.

This BWS experiment design was constructed using an orthogonal main effect plan (OMEP) in R (R Core Team, 2017), under the assumption that attributes had negligible correlations. The design was balanced (i.e., all attribute levels appear the same number of times) and orthogonal (i.e., each attribute level can be evaluated independently of other attribute levels). The design resulted in 36 profiles, each resembling a set of attribute descriptions of a hypothetical health care facility. These 36 profiles were then divided into two versions, each containing 18 choice tasks. An example of a final BWS task is shown in Figure 1 (its original version in the Appendix C).

**Fig.1 Example of a health provider profile seen in the choice tasks (translated version)**



A special BWS choice task was inserted in the questionnaire to test the rationality of responses. It only contains three attribute levels: 20 minutes, 40 minutes, and 60 minutes of travel time from home to hospital. Respondents were expected to select the first one as the best feature and the last one as the worst feature. Those who failed to do this were excluded from our final analysis.

In addition to the BWS tasks, respondents’ characteristics were also collected. A pilot study was conducted among 25 adult residents in Shanghai in June 2017 to test the validity of the questionnaire. Modifications on language were undertaken to make the questionnaire more understandable, based on feedback from the pilot.

## *2.3 Data collection*

The data were collected in Shanghai during July to October 2017. We recruited participants in two communities, two community clinics, and two tertiary hospitals in Shanghai. The two communities are in Xuhui District and Putuo District. We chose a clinic in each community. These two communities were chosen based on their GDP per capita. Shanghai Statistics Yearbook 2017 suggests that the GDP per capita of Xuhui ranks the 5th among 16 districts in Shanghai while that of Putuo ranks the 12th (Shanghai Bureau of Statistics, 2017). Sampling from these two districts therefore covered areas with different socioeconomic development levels. The two tertiary hospitals chosen are Shanghai Sixth People’s Hospital in Xuhui and the Central Hospital of Putuo. We chose these two hospitals because they primarily attract local patients, in contrast to more famous tertiary hospitals where large portions of patients are not from Shanghai.

Shanghai residents who were at least 18 years old were eligible to participate in this study. The patients recruited at community clinics and tertiary hospitals were selected if their visit was for mild disease treatment. A convenience sampling strategy was adopted. We first approached the management level of the selected community offices, community clinics and tertiary hospitals and secured their support for our study. On the agreed dates for interviews, we sent five trained interviewers to these locations and their administrative staff helped recruit participants (who were visiting that day) and guided them to the interview room one by one where our interviewer was waiting. The interviewer explained the purpose of the survey to the respondent, the choice context, and the definition of attributes. Written informed consent was obtained from all participants prior to the interview. The Shanghai Jiao Tong University Ethics Committee at School of International and Public Affairs approved this study.

As each respondent makes many choices (36 choices in our case), typically a relatively small sample size is required for BWS[[5]](#footnote-5) (Cheung et al., 2016). Following the rule of thumb proposed by Johnson and Orme (2003) and Orme (1998), we settled on a minimal sample size as 39. Those who failed the rationality test and/or did not complete the survey were excluded from data analysis. The interviewers used the two versions of questionnaire in turn during interviews, to ensure that the sample size between versions is balanced.

In the data collection we focused on the community residents and targeted a sample size of around 100. While respondents were obtained using a convenience sampling approach, we considered two districts with different socioeconomic development levels and interview dates were randomly chosen. We therefore expected the sample to be representative of the adult population in Shanghai to some extent. Regarding the two patient samples collected from tertiary hospitals and community clinics, we targeted a sample size of around 50 each. It should be noted that the three samples we collected are conceptually different and thus should not be pooled for analysis. They are expected to be composed of different types of people and present different preferences. The two patient samples already revealed their choices so are used as comparators to our main sample, allowing us to gain insights on what types of people (in terms of their characteristics and preferences) would choose tertiary hospitals or community clinics and what needs to be done to change the former group’s choice.

## *2.4 Econometric analysis*

The marginal sequential approach (SEQ) was adopted to model the responses. While the BWS data can be modelled in several ways (A. Marley, Flynn, & Louviere, 2008; A. A. Marley & Louviere, 2005), the SEQ approach has become the preferred one (Louviere et al., 2015). It assumes that respondents choose the best feature first and then the worst from the remaining ones. Therefore, each BWS task generates two sequential choice sets: (1) “choosing the best from features” and (2) “choosing the worst from features” where the best feature in the first choice set is removed. The data can then be organized based on these two choice sets and analyzed using the conditional logit or the mixed logit (MIXL) model with the latter accommodating the preference heterogeneity (McFadden & Train, 2000).

Of course, the data is slightly different from the typical choice data where the choice is always the preferred option (such as DCE). Here we dummy coded all the attribute levels with “few tests/examinations available” selected as the reference level[[6]](#footnote-6) and coded the attribute levels in the “best” choice set using 0 and 1, and those in the “worst” choice set using 0 and -1. This essentially assumes that the value of attribute levels in the “worst” choice set to be the negative of their value in the “best” choice set.

Under MIXL, the utility that respondent derives from choosing feature in choice set is given by where represents the dummy coded th attribute level and is a random coefficient assumed to be normally distributed. The errors are independently and identically distributed as type 1 extreme value. All the random coefficients are assumed to be independent with each other.

STATA 14.2 (StataCorp. 2015) was used for data analysis. The MIXL model was estimated by simulated maximum likelihood using the STATA command developed by Hole (2007). There are a relatively large number of random coefficients in our model. Hence, we generated our own Scrambled Halton draws for simulation, which can outperform the standard Halton draws when the dimensionality of integration is high (Bhat, 2003). This approach needs much less Scrambled Halton draws to achieve the same level of precision. Furthermore, to ensure the reliability of the parameter estimates, we estimated the model many times by increasing the number of draws by 500, starting from 500 draws. Estimation stability was achieved upon using 4000 draws, which led to our final estimates.

The model parameter represents the value that respondents placed on an attribute level relative to the reference level. This allows us to obtain a full ranking of all attribute levels based on the values attached to them. The model parameter does not directly reflect the preference weight within an attribute (i.e., the marginal utility obtained from moving from the reference level to another level within an attribute, e.g., physician type changing from GP to specialist) because a common reference level is used for all the attribute levels. However, it is straightforward to derive the preference weights by calculating the difference between two random coefficients. For example, the preference weight for physician type changing from GP to specialist is obtained by computing which still follows a normal distribution[[7]](#footnote-7). The delta method was used for the calculation.

A byproduct of calculating the preference weights within attributes is the relative importance across attributes (i.e., which attribute is more important overall). For example, if the mean of is larger than the mean of , this suggests respondents on average care more about physician type than cost when choosing a health care facility. A more formal approach to derive the relative importance across attributes is the rescaling method outlined in Krucien, Watson, and Ryan (2017). It normalizes the relative importance scores on a 1-100 scale.

A simulation exercise was undertaken to evaluate how to direct the patient flow from tertiary hospitals to community clinics. We constructed nine hypothetical scenarios for community residents. Under the base scenario the tertiary hospital was described as a highly professional health care institute with poor service – long waiting time, rude and cold doctors, and repeated tests and examinations. This reflects general public’s opinion on tertiary hospitals (Barber, Borowitz, Bekedam et al., 2014). The community clinic was also described based on the current situation. We created seven scenarios, in which the tertiary hospital always remained the same while the community clinic differed from the base scenario only by one attribute. In the eighth scenario, we modified all the attributes of the community clinic to an ideal level and compared it with the tertiary hospital. We then estimated the probability of choosing the community clinic over the tertiary hospital in each scenario. This allowed us to explore how changing a level among community clinic attributes would impact on its uptake. This simulation exercise was repeated for the patient sample who already chose tertiary hospitals[[8]](#footnote-8), which is probably more relevant to our key question.

The predicted probability is random given the model parameters are random. In this case, 10,000 draws were sampled from its distributions and then the mean and standard error were estimated using these draws. Furthermore, we also accounted for parameter uncertainty by using all information in the parameter distribution, including the covariance matrix rather than just their mean and standard deviation. As Hensher and Greene (2003) noted, this is preferred since using just the mean and standard deviation ignores the sampling variance in the point estimates.

# **Results**

*3.1 Respondent characteristics*

A total of 220 residents and patients participated in the survey. We excluded the responses of 18 participants because they failed in the rationality test (N=3) or provided an incomplete response (N=15). Among the 18 excluded, two community residents failed the rationality test and three did not complete the questionnaire; five hospital patients did not provide complete response while no one failed the rationality test; and one community clinic patient did not pass the rationality test and seven provided incomplete responses. Finally, 202 subjects were included in the analysis, including 97 community residents, 56 tertiary hospital patients, and 49 clinic patients. Table 2 summarizes some of their characteristics, in contrast to the population statistics.

**Table 2. Sample statistics of community residents, tertiary hospital patients, and community clinic patients.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Characteristics | Community residents  (N=97) | | Hospital patients  (N=56) | | Clinic patients  (N=49) | | General population in Shanghai \* |
|  | N | % | N | % | N | % | % |
| *Age* |  |  |  |  |  |  |  |
| 18-24 years | 11 | (11.3%) | 8 | (14.3%) | 1 | (2.0%) | 18.7% |
| 25-59 years | 57 | (58.8%) | 43 | (76.8%) | 24 | (49.0%) | 49.7% |
| 60 years or above | 29 | (29.9%) | 5 | (8.9%) | 24 | (49.0%) | 31.6% |
|  |  |  |  |  |  |  |  |
| *Sex* |  |  |  |  |  |  |  |
| Male | 39 | (40.2%) | 28 | (50.0%) | 24 | (49.0%) | 49.6% |
| Female | 58 | (59.8%) | 28 | (50.0%) | 25 | (51.0%) | 50.4% |
|  | |  |  |  |  |  |  |
| *Education ¶* | |  |  |  |  |  |  |
| Middle school or lower | 34 | (35.1%) | 12 | (21.4%) | 23 | (46.9%) | N/A |
| Secondary school | 27 | (27.8%) | 15 | (26.8%) | 20 | (40.8%) | N/A |
| Post-secondary | 36 | (37.1%) | 29 | (51.8%) | 6 | (12.2%) | N/A |
|  | |  |  |  |  |  |  |
| *Annual income§* | |  |  |  |  |  |  |
| Low income | 14 | (14.3%) | 8 | (14.7%) | 5 | (10.0%) | 15.1% |
| Middle income | 59 | (60.6%) | 35 | (62.1%) | 34 | (69.4%) | 63.9% |
| High income | 24 | (25.1%) | 13 | (23.2%) | 10 | (19.7%) | 21.0% |
|  | |  |  |  |  |  |  |
| *Employment* | |  |  |  |  |  |  |
| Employed | 66 | (68.0%) | 46 | (82.1%) | 27 | (55.1%) | 73.5% |
| Unemployed | 2 | (2.1%) | 2 | (3.6%) | 1 | (2.0%) | 3.9% |
| Retired | 24 | (24.7%) | 3 | (5.4%) | 21 | (42.9%) | 20.2% |
| Students | 5 | (5.2%) | 5 | (8.9%) | 0 | (0.0%) | 2.4% |
|  | |  |  |  |  |  |  |
| *Self-report health* | |  |  |  |  |  |  |
| Very poor | 8 | (8.2%) | 6 | (10.7%) | 6 | (12.2%) | N/A |
| Poor | 31 | (32.0%) | 9 | (16.1%) | 18 | (36.7%) | N/A |
| Average | 27 | (27.8%) | 21 | (37.5%) | 16 | (32.7%) | N/A |
| Good | 26 | (26.8%) | 17 | (30.4%) | 3 | (6.1%) | N/A |
| Excellent | 5 | (5.2%) | 3 | (5.4%) | 6 | (12.2%) | N/A |

Note:

\* The characteristics information, except income, of the general population in Shanghai was retrieved from Shanghai Statistics Yearbook 2017 (Shanghai Bureau of Statistics, 2017).

¶ Up-to-date educational distribution is unavailable for the general population in Shanghai.

§ Income groups (RMB): Low: <22000, Middle: 22001-80000, High: >80000. The statistics of income of the general population in Shanghai was retrieved from Huang (2019).

We undertook chi-square tests for age, sex, income, and employment status between the community resident sample and the population. In all cases the null hypothesis could not be rejected, suggesting that the sample is representative of the target population based on age, sex, income, and employment status.[[9]](#footnote-9)

Apart from sex and income, three samples differ across all the other characteristics. Those who chose community clinics are the oldest, least educated, mostly retired and worst in self-reported health, while those who chose tertiary hospitals appear to be the youngest, most educated, mostly employed and best in self-reported health. The community resident sample’s characteristics sit in the middle, which is sensible since it likely includes people who would use tertiary hospitals and people who would use clinics.

*3.2 Model estimates and ranking of attribute levels*

Table 3 presents the results of the MIXL model for all three samples.[[10]](#footnote-10) Most of the mean values are statistically significant, revealing clear cardinality of preference strength against the reference. Most of the standard deviation estimates are also statistically significant, suggesting the existence of strong preference heterogeneity. For the community resident sample, relative to the level of “few tests/examinations available”, participants placed positive mean values with most attribute levels except two: “5 hours of total visit time” and “being cold or rude to patients”. The positive values placed on those attribute levels indicated that respondents found them more desirable than “few tests/examinations available”. The negative values indicated the opposite.

**Table 3.** **Mixed Logit model estimates and ranking of attribute levels of community residents, tertiary hospital patients, and community clinic patients.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Attribute | Level | Community residents | | |  | Hospital patients | | |  | Clinic patients | | |
| Mean | SD | Rank |  | Mean | SD | Rank |  | Mean | SD | Rank |
| Cost | CNY 100 | 2.875\*\*\* | 0.564\*\*\* | 5 |  | 5.788\*\*\* | 0.448\* | 2 |  | 4.449\*\*\* | 1.942\*\*\* | 1 |
|  | [0.142] | [0.181] |  |  | [0.261] | [0.271] |  |  | [0.298] | [0.276] |  |
| CNY 325 | 2.285\*\*\* | 0.516\*\*\* | 7 |  | 3.572\*\*\* | 1.729\*\*\* | 5 |  | 2.669\*\*\* | 1.250\*\*\* | 5 |
|  | [0.147] | [0.146] |  |  | [0.288] | [0.172] |  |  | [0.255] | [0.243] |  |
| CNY 500 | 0.608\*\*\* | 1.321\*\*\* | 13 |  | 0.426 | 1.981\*\*\* | 12 |  | -0.282 | 1.358\*\*\* | 13 |
|  | [0.175] | [0.149] |  |  | [0.301] | [0.291] |  |  | [0.245] | [0.174] |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Type of physician | GP | 2.125\*\*\* | 0.978\*\*\* | 9 |  | 2.458\*\*\* | 1.767\*\*\* | 9 |  | 3.412\*\*\* | 1.122\*\*\* | 2 |
|  | [0.171] | [0.135] |  |  | [0.244] | [0.255] |  |  | [0.227] | [0.162] |  |
| Specialist | 3.060\*\*\* | 1.060\*\*\* | 2 |  | 7.551\*\*\* | 1.804\*\*\* | 1 |  | 2.823\*\*\* | 0.918\*\*\* | 3 |
|  | [0.164] | [0.137] |  |  | [0.349] | [0.319] |  |  | [0.222] | [0.153] |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Tests/examinations availability | Few (ref) | ─ | ─ | 15 |  | ─ | ─ | 15 |  | ─ | ─ | 11 |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Many | 3.287\*\*\* | 1.229\*\*\* | 1 |  | 4.232\*\*\* | 1.004\*\*\* | 3 |  | 2.725\*\*\* | 1.686\*\*\* | 4 |
|  | [0.154] | [0.131] |  |  | [0.256] | [0.211] |  |  | [0.239] | [0.227] |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Variety of medications | Limited | 0.142 | 0.629\*\*\* | 14 |  | 0.138 | 1.084\*\*\* | 14 |  | -0.401\*\* | 0.885\*\*\* | 14 |
|  | [0.118] | [0.132] |  |  | [0.218] | [0.156] |  |  | [0.197] | [0.164] |  |
| Great | 2.609\*\*\* | 0.691\*\*\* | 6 |  | 3.597\*\*\* | 1.040\*\*\* | 4 |  | 1.999\*\*\* | 0.264 | 8 |
|  | [0.132] | [0.110] |  |  | [0.241] | [0.183] |  |  | [0.204] | [0.191] |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Travel time to hospital | 20 mins | 2.928\*\*\* | 0.947\*\*\* | 4 |  | 3.260\*\*\* | 1.295\*\*\* | 7 |  | 2.141\*\*\* | 0.095 | 6 |
|  | [0.153] | [0.167] |  |  | [0.295] | [0.239] |  |  | [0.220] | [0.212] |  |
| 40 mins | 1.926\*\*\* | 0.767\*\*\* | 10 |  | 2.350\*\*\* | 0.939\*\*\* | 10 |  | 0.886\*\*\* | 0.882\*\*\* | 10 |
|  | [0.151] | [0.124] |  |  | [0.265] | [0.176] |  |  | [0.237] | [0.185] |  |
| 60 mins | 0.688\*\*\* | 0.060 | 12 |  | 0.361\* | 0.203 | 13 |  | -0.428\*\* | 0.656\*\*\* | 15 |
|  | [0.131] | [0.224] |  |  | [0.188] | [0.285] |  |  | [0.210] | [0.201] |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Total visit time in hospital | 1 hour | 2.196\*\*\* | 1.174\*\*\* | 8 |  | 3.567\*\*\* | 2.432\*\*\* | 6 |  | 1.053\*\*\* | 1.241\*\*\* | 9 |
|  | [0.169] | [0.127] |  |  | [0.290] | [0.248] |  |  | [0.233] | [0.226] |  |
| 3 hours | 0.836\*\*\* | 0.964\*\*\* | 11 |  | 1.131\*\*\* | 1.611\*\*\* | 11 |  | -0.196 | 0.806\*\*\* | 12 |
|  | [0.152] | [0.155] |  |  | [0.274] | [0.223] |  |  | [0.211] | [0.221] |  |
| 5 hours | -0.680\*\*\* | 0.823\*\*\* | 16 |  | -1.527\*\*\* | 1.228\*\*\* | 17 |  | -1.755\*\*\* | 0.310 | 17 |
|  | [0.137] | [0.128] |  |  | [0.222] | [0.316] |  |  | [0.182] | [0.231] |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Friendliness of doctors | No | -0.868\*\*\* | 1.580\*\*\* | 17 |  | -1.333\*\*\* | 1.561\*\*\* | 16 |  | -1.500\*\*\* | 1.874\*\*\* | 16 |
|  | [0.140] | [0.136] |  |  | [0.205] | [0.263] |  |  | [0.219] | [0.255] |  |
| Yes | 2.935\*\*\* | 1.666\*\*\* | 3 |  | 2.904\*\*\* | 1.815\*\*\* | 8 |  | 2.035\*\*\* | 1.807\*\*\* | 7 |
|  | [0.155] | [0.153] |  |  | [0.266] | [0.170] |  |  | [0.242] | [0.154] |  |

Note:

1. SD: standard deviation. Ref: reference. GP: general practitioner. CNY: Chinese Yuan. No: cold or rude. Yes: friendly or considerate.

2. [ ]: standard error. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

As BWS directly compares attribute levels, we derived a full ranking of all attribute levels. For community residents, the three top-ranked levels include “many tests/examinations”, “specialist”, and “friendly or considerate”. Those who chose tertiary hospitals ranked “specialist”, “CNY 100”, and “many tests/examinations” as their top three. This is an interesting finding, which suggests that they desire the best quality for the cheapest price. By contrast, those who chose community clinics ranked “CNY 100” as their top suggesting they are also sensitive to price. Somewhat unexpectedly, they ranked “GP” and “specialist” as the next two on top and slightly favored the former. This suggests they value the quality of doctors, but also believe GPs are more appropriate for simple care treatment.

All three samples placed “doctors being cold or rude” and “5 hours of total visit time” at the bottom of the rank order. The other bottom three factor is “many tests/examinations” for community residents and tertiary hospital patients and “60 minutes travel time” for community clinic patients, suggesting one reason why people chose local community clinics is to avoid long travel times.

*3.3 Preference weights*

Table 4 presents the preference weights for levels within each attribute, derived from the MIXL estimates. All the weights are statistically significant and with expected direction. For community residents, the attribute level of doctors being friendly/considerate has the largest preference weight 3.803 (relative to doctors being cold or rude). This is confirmed by Figure 2 which compares attributes’ impact on a common 0-100 scale. Their relative importance by decreasing order is friendliness of doctors, tests/examination availability, total visit time, variety of medicines available, cost, travel time, and physician type. The physician type turned out to be the least important attribute.

Those who chose tertiary hospitals placed large preference weight on “specialist” relative to “GP” whilst those who chose community clinics placed a small preference weight on physician type. Consequently, the physician type attribute is one of the most important attributes for the former but the least important for the latter. This sharp contrast is illustrated in Figure 2. The same situation happens to the attribute “total visit time” which the two groups valued differently.

**Table 4. Preference weights converted from the Mixed Logit estimates**

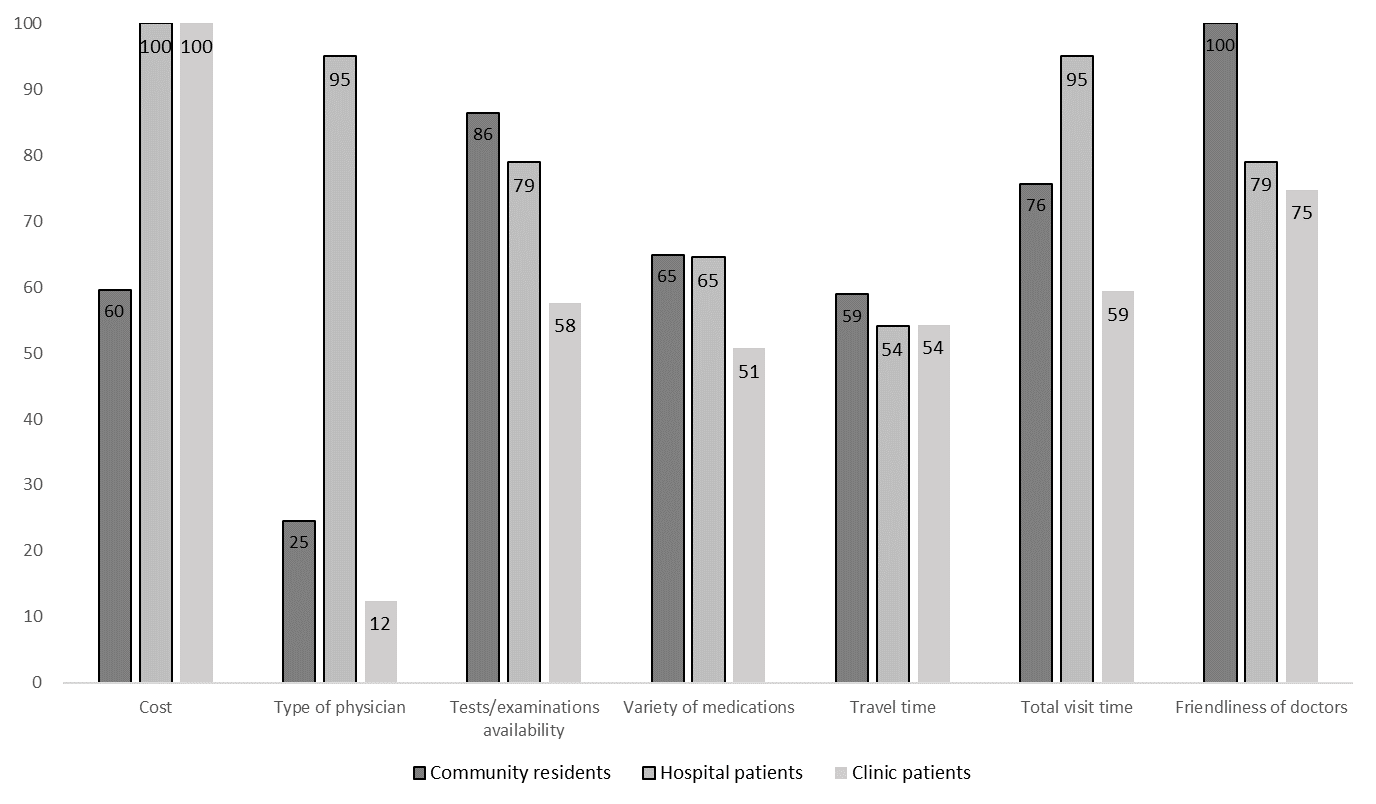
|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Attribute | Level | Community residents | |  | Hospital patients | |  | Clinic patients | |
| Mean | SD |  | Mean | SD |  | Mean | SD |
| Cost | CNY 100 | 2.267\*\*\* | 1.436\*\*\* |  | 5.362\*\*\* | 2.031\*\*\* |  | 4.730\*\*\* | 2.370\*\*\* |
|  | [0.195] | [0.162] |  | [0.357] | [0.297] |  | [0.346] | [0.255] |
| CNY 325 | 1.677\*\*\* | 1.418\*\*\* |  | 3.146\*\*\* | 2.630\*\*\* |  | 2.951\*\*\* | 1.846\*\*\* |
|  | [0.198] | [0.155] |  | [0.373] | [0.255] |  | [0.288] | [0.213] |
| CNY 500 (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
|  |  |  |  |  |  |  |  |  |  |
| Type of physician | GP (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
| Specialist | 0.936\*\*\* | 1.442\*\*\* |  | 5.093\*\*\* | 2.525\*\*\* |  | -0.589\*\* | 1.450\*\*\* |
|  | [0.201] | [0.136] |  | [ 0.365] | [0.296] |  | [0.235] | [0.155] |
|  |  |  |  |  |  |  |  |  |  |
| Tests/examinations availability | Few (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
| Many | 3.287\*\*\* | 1.229\*\*\* |  | 4.232\*\*\* | 1.004\*\*\* |  | 2.725\*\*\* | 1.686\*\*\* |
|  | [0.154] | [0.131] |  | [0.256] | [0.211] |  | [0.239] | [0.227] |
|  |  |  |  |  |  |  |  |  |  |
| Variety of medications | Limited (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
| Great | 2.467\*\*\* | 0.934\*\*\* |  | 3.459\*\*\* | 1.503\*\*\* |  | 2.400\*\*\* | 0.923\*\*\* |
|  | [0.143] | [0.125] |  | [0.273] | [0.181] |  | [0.223] | [0.168] |
|  |  |  |  |  |  |  |  |  |  |
| Travel time to hospital | 20 mins | 2.240\*\*\* | 0.949\*\*\* |  | 2.899\*\*\* | 1.311\*\*\* |  | 2.568\*\*\* | 0.663\*\*\* |
|  | [0.171] | [0.167] |  | [0.317] | [0.232] |  | [0.248] | [0.197] |
| 40 mins | 1.238\*\*\* | 0.770\*\*\* |  | 1.990\*\*\* | 0.961\*\*\* |  | 1.313\*\*\* | 1.099\*\*\* |
|  | [0.171] | [0.124] |  | [0.290] | [0.183] |  | [0.276] | [0.191] |
| 60 mins (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
|  |  |  |  |  |  |  |  |  |  |
| Total visit time in hospital | 1 hour | 2.876\*\*\* | 1.434\*\*\* |  | 5.093\*\*\* | 2.725\*\*\* |  | 2.807\*\*\* | 1.279\*\*\* |
|  | [0.192] | [0.128] |  | [0.325] | [0.243] |  | [0.241] | [0.232] |
| 3 hours | 1.516\*\*\* | 1.267\*\*\* |  | 2.657\*\*\* | 2.026\*\*\* |  | 1.558\*\*\* | 0.864\*\*\* |
|  | [0.179] | [0.150] |  | [0.333] | [0.260] |  | [0.221] | [0.211] |
| 5 hours (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
|  |  |  |  |  |  |  |  |  |  |
| Friendliness of doctors | No (ref) | ─ | ─ |  | ─ | ─ |  | ─ | ─ |
| Yes | 3.803\*\*\* | 2.296\*\*\* |  | 4.237\*\*\* | 2.394\*\*\* |  | 3.535\*\*\* | 2.604\*\*\* |
|  | [0.178] | [0.143] |  | [0.323] | [0.209] |  | [0.291] | [0.223] |

Note:

1. SD: standard deviation. Ref: reference. GP: general practitioner. CNY: Chinese Yuan. No: cold or rude. Yes: friendly or considerate.

2. [ ]: standard error. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. All estimates have p<0.01 except for the mean estimate of “Specialist” for clinic patients which has p<0.05.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

****

**Fig.2 Relative importance of attributes for community residents, hospital patients, and clinic patients.**

*3.4 Simulation results*

Table 5 and 6 present the results of the simulation exercises where we predicted probability of choosing a community clinic over a tertiary hospital in nine hypothetical scenarios for community residents and hospital patients, respectively. Table 5 suggests that, in the base scenario (which intends to reflect the current situation), only 26% of community residents would choose the community clinic over tertiary hospital. When the cost of services was decreased to CNY 100 while the other attributes remained the same as in the base case (scenario 1), the proportion of residents choosing the community clinic increased to 31%. When the physicians in the community clinic become specialists (scenario 2), the proportion increased to 33%. When the number of tests and examinations increased (scenario 3), the proportion increased to 62%. When there were more medications available (scenario 4), the proportion increased to 52%. A shorter traveling time from home to community clinic (20 minutes in scenario 5) increased the proportion to 36%. When the total visit time was decreased from 3 hours to 1 hour (scenario 6), the proportion increased to 40%. The change on friendliness of doctors from being cold or rude to friendly and considerate (scenario 7) increased the proportion to 64%. We noticed that, among all the single measures improving the attractiveness of community clinics, the upgrade on tests/examinations availability or friendlies of doctors increased the proportions of residents who were willing to visit community clinic from less than 30% to more than 60%. Finally, when we set all the attribute at the most attractive level (scenario 8), almost all residents would prefer the community clinic to tertiary hospital.

Based on Table 6, only 16% of tertiary hospital patients would choose the community clinic over the tertiary hospital in the base scenario. The rest of the analysis is similar to above, but the most striking impact happens in scenario 2 where specialists replace GPs in the community clinic which more than tripled its choosers. The other two largest impacts are in scenario 3 and 7, where improvement on the availability of tests and examinations and doctors’ attitude increased the chance of choosing community clinics to 44% in each case. This shows that these changes are the most effective among all the single measures in directing hospital patient flow to community clinics.

**Table 5. Uptake predictions in hypothetical scenarios: community residents**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Tertiary hospital | Community clinic | | | | | | | | |
|  |  | Base scenario | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 | Scenario 5 | Scenario 6 | Scenario 7 | Scenario 8 |
| Cost (CNY) | 500 | 325 | 100 | 325 | 325 | 325 | 325 | 325 | 325 | 100 |
| Type of physician | specialist | GP | GP | specialist | GP | GP | GP | GP | GP | specialist |
| Tests/examinations availability | many | few | few | few | many | few | few | few | few | many |
| Variety of medications | great | limited | limited | limited | limited | great | limited | limited | limited | great |
| Travel time to hospital (minutes) | 60 | 40 | 40 | 40 | 40 | 40 | 20 | 40 | 40 | 20 |
| Total visit time in hospital (hours) | 5 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 3 | 1 |
| Friendliness of doctors | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | friendly or considerate | friendly or considerate |
| Predicted probability of choosing community clinic [SE] |  | 0.255 [0.040] | 0.313 [0.045] | 0.333 [0.043] | 0.620 [0.046] | 0.520 [0.049] | 0.357 [0.046] | 0.395 [0.047] | 0.641  [0.043] | 0.998  [0.001] |

**Table 6. Uptake predictions in hypothetical scenarios: hospital patients**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Tertiary hospital | Community clinic | | | | | | | | |
|  |  | Base scenario | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 | Scenario 5 | Scenario 6 | Scenario 7 | Scenario 8 |
| Cost (CNY) | 500 | 325 | 100 | 325 | 325 | 325 | 325 | 325 | 325 | 100 |
| Type of physician | specialist | GP | GP | specialist | GP | GP | GP | GP | GP | specialist |
| Tests/examinations availability | many | few | few | few | many | few | few | few | few | many |
| Variety of medications | great | limited | limited | limited | limited | great | limited | limited | limited | great |
| Travel time to hospital (minutes) | 60 | 40 | 40 | 40 | 40 | 40 | 20 | 40 | 40 | 20 |
| Total visit time in hospital (hours) | 5 | 3 | 3 | 3 | 3 | 3 | 3 | 1 | 3 | 1 |
| Friendliness of doctors | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | cold or rude | friendly or considerate | friendly or considerate |
| Predicted probability of choosing community clinic [SE] |  | 0.160 (0.036) | 0.278 (0.051) | 0.504 (0.060) | 0.435 (0.055) | 0.371 (0.056) | 0.211 (0.041) | 0.313 (0.044) | 0.443  (0.064) | 0.999  (0.0002) |

# **Discussion**

Underuse of primary care facilities and overcrowding of tertiary hospitals have become one of the major challenges the Chinese healthcare system is facing today. The multi-tiered system has failed to function as intended. Substantial resources have been wasted and the deteriorating relationship between patients and doctors in tertiary hospitals has been in the spotlight of society. As Wu and Lam (2016) pointed out, many factors have caused this problem and reforms at different dimensions are required for a fundamental solution. In this study, however, we investigate this problem by focusing on examining patients’ choice of healthcare facilities. With the absence of a gate-keeping mechanism, understanding how patients choose their care providers might be the most pragmatic way to generate useful insights for redistributing patient flows from tertiary hospitals to clinics.

A BWS experiment was undertaken in Shanghai as an example of China’s urban areas. Three samples were obtained, including community residents who are representative of the Shanghai adult population, patients at tertiary hospitals and patients at community clinics. Results suggest that these groups are composed of different people and possess some different preference patterns. For community residents, sufficient tests and examinations, doctors with expertise, and good service attitude are on average the most desirable features of a healthcare facility, while unfriendly practitioners and lengthy visit are the most unattractive features. Regarding the attributes, friendliness of doctors and tests and examination availability are the two major drivers of patient choice. Travel time and total visit time are also important factors, but since they are the strength of the community clinics in Shanghai, we discuss the policy implications of the other attributes’ results.

The type of physician is described using two levels: GP and specialist. One thing to be kept in mind is that stated preference methods such as BWS simplify the real-world attributes using a small number of levels. Here GP and specialist were selected to represent different levels of medical training experience and expertise, or the quality of physicians, which can be vastly different across healthcare facilities. From the perspective of community residents, specialist is ranked at No. 2 and GP at No.9 among all 17 attribute levels. However, similar values are attached to both levels, suggesting that residents would consider a facility operated by specialists highly desirable but they also value one operated by GPs.

The estimates using the tertiary hospital patient sample suggested a somewhat different story. Specialist and GP were ranked at No. 1 and No. 9, similar to the community residents’ case. However, the values attached to the two levels are much larger, implying they found a facility operated by specialists much more attractive than one operated by GPs. The physician type was one of the most important factors and drove these patients to choose tertiary hospitals. This also means, if the government wants to direct people from tertiary hospitals to community clinics, it needs to ensure the physician delivers good quality care that can be recognized as such by patients.

This is consistent with the empirical evidence from the literature (Yang & Yang, 2009). China has a multi-tiered medical education system, with tier 1 providing degree-oriented medical training, tier 2 providing 3-year vocational training for high school graduates, and tier 3 providing limited vocational training for middle school graduates. This education system intends to trade-off between quantity and quality in the hope to meet the rapidly increasing healthcare demand. However, this medical education stratification of doctors has led to the aggregation of doctors with different education levels at different care levels, resulting in under supply of GPs who received sufficient training to community clinics (Hsieh & Tang, 2019; Xian, Lu, & Lv, 2012). In many community clinics, GPs are junior physicians or those without sufficient medical training. These have led to people’s distrust of the professionalism of doctors in community clinics, which has become the major obstacle for patients to choose clinics as their first choice for healthcare services. Establishing the trust of the quality of doctors by the population is therefore a crucial step by the government to direct the patient flow from hospitals to community clinics.

Physicians being cold or rude was the least desirable feature for all three groups considered in the study. For the community residents, physicians being friendly or considerate was also ranked at No.3. Due to this drastic difference between the values attached to these two levels, the friendliness of doctors generated the highest relative importance score among all attributes and the largest impact of the uptake on community clinics in the simulation. This result reflects the ever-increasing tension between patients and physicians in hospitals, ironically the exact consequence of the failure of distributing patients based on their actual medical needs (Tucker, Cheng, Wong et al., 2015). Nevertheless, our simulation suggests that improving a doctors’ attitude towards patients at community clinics might be the most efficient way to attract more patients in the current multi-tiered health system. This in turn may break the spiral of overcrowding tertiary hospitals and worsening patient-physician relationship.

Low service cost (CNY 100) was one of the most desirable features of a healthcare facility, in particular for the two patient samples collected at the tertiary hospitals and community clinics. Indeed, the financial burden of care services has been complained about by patients for long (Tang, Tao, & Bekedam, 2012). The Chinese government has started taking action to reduce the cost, such as the drug mark-up removal conducted across China in 2017 (Zang, Zhang, Wei et al., 2019) and the collective pharmaceutical procurement (Jiang, Chen, Wu et al., 2020). Measures in reducing the cost of primary care services and improving reimbursement policy for visiting community clinics should be further explored and implemented.

The service capacity of healthcare facilities is represented by the availability of accurate tests/examinations using medical equipment and the availability of effective medications. They were both identified as major drivers of patient choice, with the former rated more important. For a long period of time, community clinics have lacked financial support from their local municipal fiscal system and unsurprisingly been perceived as places with poor service capacity (Wang, Gusmano, & Cao, 2011). However, the situation is currently improving with the government investing substantial resources in expanding their capacity. In the meantime, regulations on medication provisions should be reviewed and updated. The variety of drugs in community clinics is currently regulated by National Essential Drug List (Song, Bian, Petzold et al., 2014), which has been found to be unable to satisfy urban patients’ needs for primary care (Zhang, Yang, & Li, 2017). A drug list with greater variety of medications for use in community clinics is potentially warranted.

Our study has several limitations. First, the administrative staff who helped us recruit participants did not record exactly how many residents or patients they approached and therefore we cannot calculate the response rate. Their impression is that the response rate is high as only a few declined due to their time constraints. Second, while BWS has several advantages over DCE, it also has its own constraints. We were unable to use it to estimate willingness to pay for changes in the attributes which would help us derive how much dollar values are attached to each attribute. We were also unable to undertake an accurate market share prediction since BWS is a forced choice task. The policy simulation we conducted is largely one way to evaluate the relative impact of an attribute with a base case set up (Lancsar, Louviere, & Flynn, 2007). Finally, like all the stated preference methods, BWS cannot avoid hypothetical bias, i.e., respondents may act differently in real world from what they choose in our survey.

# **Conclusions**

This study adds to the literature on patient choice of healthcare facilities. In contrast to the broad literature aiming to test how patients respond to quality signals, we focused on identifying the relative importance of attributes and their levels as our primary objective was to uncover the drivers of patient’s choice under Chinese multi-tiered hospital system. Our findings offered insights into helping understand why Chinese patients prefer tertiary hospitals over community clinics even for simple care needs, and how to effectively redistribute patients from tertiary hospitals to community clinics.

References

Barber, Borowitz, Bekedam, et al. (2014). The hospital of the future in China: China’s reform of public hospitals and trends from industrialized countries. *Health Policy and Planning, 29*(3), 367-378.

Bhat. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B: Methodological, 37*(9), 837-855.

Brekke, Gravelle, Siciliani, et al. (2014). Patient choice, mobility and competition among health care providers. In *Health care provision and patient mobility* (pp. 1-26): Springer.

Cheung, Wijnen, Hollin, et al. (2016). Using best–worst scaling to investigate preferences in health care. *PharmacoEconomics, 34*(12), 1195-1209.

Coast, Flynn, Natarajan, et al. (2008). Valuing the ICECAP capability index for older people. *Social science & medicine, 67*(5), 874-882.

de Bekker‐Grob, Ryan, & Gerard. (2012). Discrete choice experiments in health economics: a review of the literature. *Health economics, 21*(2), 145-172.

Finn, & Louviere. (1992). Determining the appropriate response to evidence of public concern: the case of food safety. *Journal of Public Policy & Marketing*, 12-25.

Flynn, Louviere, Peters, et al. (2007a). Best--worst scaling: What it can do for health care research and how to do it. *J Health Econ, 26*(1), 171-189. doi:10.1016/j.jhealeco.2006.04.002

Flynn, Louviere, Peters, et al. (2007b). Best–worst scaling: what it can do for health care research and how to do it. *Journal of health economics, 26*(1), 171-189.

Hensher, & Greene. (2003). The mixed logit model: the state of practice. *Transportation, 30*(2), 133-176.

Hole. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal, 7*(3), 388-401.

Hsieh, & Tang. (2019). The multi-tiered medical education system and its influence on the health care market—China’s Flexner Report. *Human resources for health, 17*(1), 50.

Huang. (2019). The income distribution in China and its evolution trend. *Reformation and Strategy, 35*(10), 66-75. doi:10.16331/j.cnki.issn1002-736x.2019.10.009

Jiang, Chen, Wu, et al. (2020). Collective pharmaceutical procurement in China may have unintended consequences in supply and pricing. *Journal of Global Health, 10*(1).

Johnson, & Orme. (2003). Getting the most from CBC. *Sequim: Sawtooth Software Research Paper Series, Sawtooth Software*.

Krucien, Watson, & Ryan. (2017). Is best–worst scaling suitable for health state valuation? A comparison with discrete choice experiments. *Health economics, 26*(12).

Lancsar, Louviere, & Flynn. (2007). Several methods to investigate relative attribute impact in stated preference experiments. *Social science & medicine, 64*(8), 1738-1753.

Liu, Kong, & de Bekker-Grob. (2019). Public preferences for health care facilities in rural China: A discrete choice experiment. *Social Science & Medicine, 237*, 112396.

Liu, Kong, Yuan, et al. (2018). Factors influencing choice of health system access level in China: a systematic review. *PloS one, 13*(8), e0201887.

Louviere, Flynn, & Marley. (2015). *Best-worst scaling: Theory, methods and applications*: Cambridge University Press.

Marley, Flynn, & Louviere. (2008). Probabilistic models of set-dependent and attribute-level best–worst choice. *Journal of Mathematical Psychology, 52*(5), 281-296.

Marley, & Louviere. (2005). Some probabilistic models of best, worst, and best–worst choices. *Journal of Mathematical Psychology, 49*(6), 464-480.

McFadden, & Train. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447-470.

McIntosh, & Louviere. (2002). Separating weight and scale value: an exploration of best-attribute scaling in health economics. *Health Economics Study Group. Odense, Denmark*.

Orme. (1998). Sample size issues for conjoint analysis studies. *Sawthooth Software Research paper Series Squim, WA, USA: Sawthooth Software Inc*.

Ratcliffe, Couzner, Flynn, et al. (2011). Valuing child health utility 9D health states with a young adolescent sample. *Applied health economics and health policy, 9*(1), 15-27.

Severin, Schmidtke, Mühlbacher, et al. (2013). Eliciting preferences for priority setting in genetic testing: a pilot study comparing best-worst scaling and discrete-choice experiments. *European Journal of Human Genetics, 21*(11), 1202-1208.

Shanghai Bureau of Statistics. (2017). *Shanghai 2017 Statistic Yearbook*. Retrieved from <https://data.cnki.net/Area/Home/Index/D09>.

Song, Bian, Petzold, et al. (2014). The impact of China’s national essential medicine system on improving rational drug use in primary health care facilities: an empirical study in four provinces. *BMC Health Services Research, 14*(1), 507.

State Council of China. (2009). Opinions of the Central Committee of the Communist Party of China and the State Council on Deepening the Reform of the Medical and Health System. In (pp. 4-7).

Sun, Lin, Zhao, et al. (2017). Reducing waiting time and raising outpatient satisfaction in a Chinese public tertiary general hospital-an interrupted time series study. *BMC Public Health, 17*(1), 668.

Tang, Tao, & Bekedam. (2012). *Controlling cost escalation of healthcare: making universal health coverage sustainable in China.* Paper presented at the BMC Public Health.

Tucker, Cheng, Wong, et al. (2015). Patient–physician mistrust and violence against physicians in Guangdong Province, China: a qualitative study. *BMJ open, 5*(10), e008221.

Wang, Gusmano, & Cao. (2011). An evaluation of the policy on community health organizations in China: will the priority of new healthcare reform in China be a success? *Health Policy, 99*(1), 37-43.

Wu, & Lam. (2016). Underuse of primary care in China: the scale, causes, and solutions. *The Journal of the American Board of Family Medicine, 29*(2), 240-247.

Xian, Lu, & Lv. (2012). Chinese General Practitioner Training Scheme: Challenges and Strategies. *Chinese General Practice, 12*(8A), 2498-2501.

Yang, & Yang. (2009). Community health service centers in China, not always trusted by the populations they serve? *China Economic Review, 20*(4), 620-624.

Yip, Hsiao, Chen, et al. (2012). Early appraisal of China's huge and complex health-care reforms. *The Lancet, 379*(9818), 833-842.

Yu, Liu, & Huang. (2007). An Empirical Study on Accessibility and Costs of Medical Services in China: Patients' Perspective (in Chinese). *Chinese Hospitals, 11*(9), 30-33.

Zang, Zhang, Wei, et al. (2019). Impact of public hospital pricing reform on medical expenditure structure in Jiangsu, China: a synthetic control analysis. *BMC health services research, 19*(1), 512.

Zhang, Yang, & Li. (2017). Analysis of the Utilization Status of Essential Medicines in Beijing’s Community Health Institutions. *Therapeutic Innovation & Regulatory Science, 51*(1), 39-44.

Zhao, Chen, & Liu. (2009). Residents' Care-Seeking Behavior and the Utilization of Community Health Centers. *Chinese Nursing Research*(8), 2058-2059.

1. These are in urban areas. In rural areas, primary care facilities include township health centers and village clinics. [↑](#footnote-ref-1)
2. Examples of context factors include capitation/gatekeeping, freedom of service choice, salary reform on health workers and public campaign/interaction of social capital. [↑](#footnote-ref-2)
3. While we did not collect data to measure respondents’ perceived difficulty of the BWS tasks, the feedback from our interviewers suggested that the respondents in general did not find the tasks difficult to complete. [↑](#footnote-ref-3)
4. It should be noted that, Liu et al. (2019) also used the same attribute with the same definition. They identified this attribute from their own literature view and focus group study independently from ours. [↑](#footnote-ref-4)
5. For example, Ratcliffe, Couzner, Flynn et al. (2011) used a sample of 16 for five choice tasks with nine attributes in each choice task. Severin, Schmidtke, Mühlbacher et al. (2013) used a sample of 26 for 12 choice tasks with six attributes in each choice task. [↑](#footnote-ref-5)
6. Focus group discussions suggest that it has a moderate importance among all attribute levels. [↑](#footnote-ref-6)
7. The addition or difference between two independent normally distributed random variables is still normally distributed. [↑](#footnote-ref-7)
8. We chose not to undertake this simulation for the community clinic patient sample since they have already chosen the community clinics as the government wanted. [↑](#footnote-ref-8)
9. For age, the chi-square statistics is 4.502, smaller than the critical value 5.99. For sex, the statistics is 3.424, smaller than the critical value 3.84. For income, the statistics is 0.819, smaller than the critical value 5.99. For employment status (the last three levels grouped), the statistics is 1.484, smaller than the critical value 3.84. [↑](#footnote-ref-9)
10. Table 3 (except hospital patients) was reproduced by including those who failed the rationality tests which had little impact on estimates (results are documented in the Appendix D) [↑](#footnote-ref-10)