Discussion Paper Series

IZA DP No. 18384 February 2026

Health Perception Biases and Healthcare Utilization in China: A Longitudinal Approach

Peng Nie
Xi'an Jiaotong University, University of Hohenheim and IZA@LISER

Sonja Spitzer
University of Vienna and Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/ÖAW, WU)

Alfonso Sousa-Poza
University of Hohenheim and IZA@LISER

IZA NETWORK @ LISER
ISSN: 2365-9793

IZA DP No. 18384 February 2026

Health Perception Biases and Healthcare Utilization in China: A Longitudinal Approach

Abstract

Using panel data on over 17,000 adults from the China Health and Retirement Longitudinal Study (CHARLS), this paper investigates the role of health perception biases for healthcare utilization and related expenditure in China. We measure health perception biases as the difference between objective health outcomes from physical examinations and self-reported health. Leveraging the longitudinal dimension of the data, we address unobserved individual heterogeneity in the relationship between perception biases and healthcare use. We find that individuals who underestimate their health visit the doctor more often and have more hospital stays, while those who overestimate their health are less likely to use those healthcare services. Health perception biases are also strongly associated with total and out-of-pocket expenditures for both outpatient and inpatient care. Importantly, family support – especially the presence of co-resident sons – mitigates the tendency of those underestimating their health to seek more care, highlighting the role of family dynamics in healthcare decisions. Moreover, differences in China's heterogeneous health insurance schemes appear to influence how health misperception translates into healthcare spending.

JEL classification

I10, I12, I18, D83, P46

Keywords

health perception bias, overconfidence, underconfidence, healthcare utilization, China

Corresponding author

Peng Nie
peng_nie@uni-hohenheim.de

1. Introduction

Biased perception, such as overconfidence, affects a wide array of human behaviors and decision-making, including consumer choices (Caliendo & Huang, 2008; Hou et al., 2018), political behavior (Ortoleva & Snowberg, 2015), labor market outcomes (Friehe & Pannenberg, 2019; Tiefenbeck et al., 2016), financial decision-making (Mazzonna & Peracchi, 2024), and corporate investment (Malmendier & Tate, 2015). Several studies have investigated the association between health perception biases – i.e. overestimating or underestimating – and health outcomes (Murphy et al., 2017; Nie et al., 2022; Pirinsky, 2013) and health behaviors (Arni et al., 2021; Lei et al., 2021; Nie et al., 2023; Spitzer & Shaikh, 2022). It is plausible that such biases also shape whether and how people engage with healthcare services – either by discouraging necessary care or prompting excessive, and possibly unnecessary, utilization. Indeed, for higher-income countries with universal healthcare systems, it has been shown that individuals who overestimate their health are less likely to visit a doctor, while those who underestimate their health are more prone to seek medical care (Spitzer & Shaikh, 2022).

However, existing evidence remains limited in two important ways. First, findings from high-income countries with well-developed healthcare infrastructures and universal coverage may not easily translate to lower- and middle-income countries, where both healthcare systems and family dynamics differ substantially. In many developing contexts, including China, access to formal healthcare is more fragmented and family members play a much more prominent role in healthcare decision-making and support. As a result, the way health perception biases influence healthcare utilization may follow different patterns that existing studies overlook. Second, the available evidence relies on cross-sectional data, which cannot account for unobserved individual heterogeneity that may simultaneously affect both health perceptions and healthcare utilization.

In this paper, we address these gaps by examining how health perception biases influence healthcare utilization and related expenditures in China – an important and underexplored context. China is one of the fastest-aging societies globally and already has the largest aging population (Nie, Li, Zhang, et al., 2021; Zhao et al., 2014). Healthcare spending is projected to grow rapidly (Dieleman et al., 2017), alongside a rising burden of noncommunicable diseases (NCDs), particularly among older adults (Wang et al., 2011). Despite the growing prevalence of NCDs, health awareness and management remain low; for example, only 38.7% of hypertensive older adults are aware of their condition (You et al., 2018), compared to much higher rates in countries like Germany (Zhou et al., 2019). These widespread gaps in health awareness suggest that health perception biases may be pervasive and consequential for healthcare decision-making.

Our analysis also explores two key mechanisms that may shape the relationship between health misperception and healthcare utilization in China. First, family support, particularly from co-resident children, may moderate this relationship. In traditional Chinese families, children play an essential role not only as caregivers and financial supporters but also as influencers of healthcare decisions (Wen & Zhang, 2023). Co-resident children may either encourage care-seeking when parents overestimate their health or temper unnecessary care when parents underestimate it. Second, China's fragmented and heterogeneous health insurance system, composed of schemes with distinct benefit packages and reimbursement rates, is likely to condition the financial consequences of health perception biases. These schemes differ widely in coverage and cost-sharing, which may mediate how perception biases translate into actual healthcare expenditures. By examining these mechanisms, we show that both family and institutional contexts crucially shape how perception biases translate into actual healthcare utilization and spending.

Our analysis uses high-quality panel data from the China Health and Retirement Longitudinal Study (CHARLS), covering over 17,000 middle-aged and older adults. We measure health perception bias as the difference between objective health measures (clinical outcomes from physical exams) and self-reported health status. Using random effects (RE) Poisson models, we examine how these biases relate to the number of doctor visits and hospital stays, while RE two-part models allow us to assess their association with both total and out-of-pocket healthcare expenditures, thus capturing the financial dimension of biased perceptions.

Our results show that individuals who underestimate their health are more likely to visit the doctor and experience hospital stays, whereas those who overestimate their health are less likely to use these healthcare services. These biases are also closely linked to both total and out-of-pocket expenditures for outpatient and inpatient care. Importantly, the presence of co-resident sons mitigates the positive association between health underestimation and doctor visits, underscoring the moderating role of family support. Additionally, differences across China's health insurance schemes influence the extent to which perception biases translate into healthcare spending.

Our study makes several contributions to the existing literature. To the best of our knowledge, it is the first to provide evidence on the link between health perception biases and healthcare utilization in a developing or emerging country. While previous research has primarily focused on developed regions, particularly the U.S. and Europe, applying findings from these contexts to China is not necessarily straightforward. Although the relationship between health misperceptions and healthcare utilization may appear general (and linear), specific characteristics of the Chinese context could challenge such generalizations. We highlight two key aspects that distinguish China from Western settings: the role of the Chinese family and the country's unique and heterogeneous health insurance system. As we demonstrate, both factors significantly shape and moderate the impact of health misperceptions on healthcare utilization. Additionally, our study contributes to the literature by leveraging high-quality panel data, enabling us to address unobserved heterogeneity, which previous studies based on cross-sectional data could not fully account for. Finally, our dataset covers a wider range of healthcare dimensions than prior research, including detailed information on healthcare expenditures, allowing for a more comprehensive analysis of healthcare utilization and its implications for both individual financial burden and broader healthcare system sustainability.

The remainder of the paper is organized as follows: Section 2 documents the past research and highlights the contribution of this paper to this research. Section 3 describes the dataset and empirical strategy, and Section 4 reports the results. Section 5 concludes the paper with a discussion of the major findings and their implications for policy.

2. Prior literature and underlying mechanisms

2.1 Prior literature

A recently growing body of literature has examined the association between health perception biases – such as overestimating or underestimating one's health – on various health outcomes and behaviors.[1] Several studies highlight that misperceptions of health or body weight are associated with both mental health and lifestyle behaviors. For example, while some evidence suggests that overconfidence may be linked to higher levels of happiness and self-esteem (Murphy et al., 2017; Pirinsky, 2013), other studies find that overestimating health or weight is associated with lower quality of life, greater stress, and depression (Heard et al., 2017; Lim & Wang, 2013; McGraw et al., 2004; Nie et al., 2022; Park et al., 2017).

In addition to health outcomes, health perception biases have been linked to health-related behaviors, including diet, exercise, smoking, and alcohol use. For instance, individuals who overestimate their health are found to engage in fewer healthy behaviors, such as exercising or maintaining a balanced diet, and are more likely to engage in risky behaviors like excessive drinking (Arni et al., 2021; Lei et al., 2021; Lim & Wang, 2013; Nie et al., 2023; Silva et al., 2021). These findings suggest that biased self-assessments of health can have broader implications for how individuals manage their health and well-being. However, few studies have directly investigated how health perception biases affect healthcare utilization. To our knowledge, only Spitzer and Shaikh (2022) examine this relationship, showing that individuals who overestimate their health are less likely to visit a doctor, whereas those who underestimate their health are more likely to seek medical care.

Several key aspects of the existing literature on the relationships between health misperception, health, and health behaviors warrant attention. First, most studies have been conducted in developed countries, particularly the U.S. and Europe. Second, the majority of these studies rely on cross-sectional data, limiting their ability to address endogeneity concerns such as reverse causality, and unobserved heterogeneity. Third, while Spitzer and Shaikh (2022) focus on doctor visits and hospital stays as proxies for healthcare utilization, other important dimensions, such as outpatient and inpatient expenditures, remain understudied. To address these gaps, our study contributes to the

[1] To our knowledge, we only identify two studies on investigate the determinants of bodyweight misperception in developing countries (Akindele et al., 2017; Saleem et al., 2013): The first study, using a cross-section study in urban Lagos State, Nigeria, shows that women are more likely to misperceive their weight. Nonetheless, Akindele et al. (2017) do not accurately define weight underestimate, accurate, and overestimate. The latter study employs a cross-sectional data set based on undergraduate university students (aged 15-24 years old) of Karachi, Pakistan, and find that females compared to males are more likely to overestimate their weight, and misperception increases with the age of respondents (Saleem et al., 2013).

literature by conducting a comprehensive analysis using nationally representative longitudinal data from the 2011–2015 CHARLS survey to investigate the association between health misperception and healthcare utilization.

2.2 Mechanisms: Family support and health insurance system

While the existing literature has identified important associations between health perception biases and health-related behaviors, less is known about the channels through which these biases influence healthcare utilization, especially in lower- and middle-income countries. In the Chinese context, two institutional and social factors are likely to play a crucial moderating role in this relationship: family support, particularly from co-resident children, and the country's fragmented and heterogeneous health insurance system.

The role of co-resident children. The co-resident children may moderate the linkage between health misperception and healthcare utilization among older Chinese adults. Specifically, children play a pivotal role in traditional Chinese families, not only as caregivers (enhancing the ease of obtaining health care for older adults) but also as financial supporters (bolstering the fiscal capacity of older adults concerning the healthcare services they can afford) (Wen & Zhang, 2023). They also influence healthcare decision-making (Wen & Zhang, 2023), acting as intermediaries in the process. When individuals overestimate their health, children may encourage them to seek medical attention regardless. Conversely, when individuals underestimate their health, children can provide a "reality check," reassuring them that medical care may not be necessary.

Examining this mechanism in the Chinese context is both significant and relevant for several reasons. First, unlike in developed countries, many developing and emerging Asian nations, particularly China, continue to uphold traditional family norms such as filial piety while also facing limitations in public pension and social security systems (Yu et al., 2022). For centuries, families have served as the primary source of old-age support, and they remain the dominant providers of both economic and functional assistance for older adults (Nie & Zhao, 2023). Co-residency plays a crucial role in this support system; in 2018, approximately half of Chinese individuals aged 60 and older lived with their children or grandchildren (Nie & Zhao, 2023), and around 67% of middle-aged and older parents maintained frequent contact with their adult children (at least once per week) (Zhao et al., 2021). Second, in contrast to many developed countries where long-term care includes both formal social care (provided by the state and private enterprises) and informal care (provided by family members), China continues to rely heavily on informal caregiving to meet the daily living needs of older adults. The demand for informal care is expected to rise significantly, from 41.3 million individuals in 2015 to 82.6 million by 2035 (Hu, 2019). Given these dynamics, understanding the role of family support in shaping healthcare decisions within the Chinese context is both timely and essential. Based on these observations, we hypothesize that the association between health misperception and healthcare utilization is moderated by co-resident children.

Heterogeneous medical insurance system. China's national medical insurance system consists of three primary schemes: Urban Employees Basic Medical Insurance (UEBMI), Urban Residents' Basic Medical Insurance (URBMI), and the New Rural Cooperative Medical Insurance Scheme (NRCMI).[2] UEBMI, established in 1998, is a mandatory insurance for urban employees and retirees (Zhang et al., 2017). Employees contribute around 2% of their salary, while employers pay 6-10%. After meeting deductibles, employees cover 10-30% of medical costs, with an annual cap limiting out-of-pocket expenses. URBMI, designed for unemployed urban residents, is voluntary and features lower premiums. The government provides substantial subsidies, and enrollees pay 30-50% of medical costs after deductibles (Huang & Gan, 2017). NRCMI, introduced in 2003 for rural residents, also operates on voluntary enrollment, with lower premiums but higher out-of-pocket costs (40-60% after deductibles).

[2] China has increasingly consolidated two of the original three schemes, merging NRCMI and URBMI into the Urban and Rural Residences' Medical Insurance (URRMI) since 2016. We, however, focus on three primary schemes because our study period is 2011-2015, which is before the implementation of URRMI.

Government subsidies play a crucial role in both URBMI and NRCMI.

Regarding healthcare expenditures, URBMI does not cover outpatient services, however, NRCMI partially covers outpatient expenditures (Nie and Zhao, 2023). Consequently, relative to NRCMI, URBMI may decrease outpatient utilization. Therefore, due to the huge discrepancy in benefit package, reimbursement rate, and cost-sharing across China's health insurance schemes we assume that different forms of healthcare insurance will mediate the effect of health misperception on healthcare expenditures.

3. Data and methods

3.1 The China Health and Retirement Longitudinal Study (CHARLS)

Our analysis is based on CHARLS, a nationally representative longitudinal survey of adults aged 45+ in China (Zhao et al., 2014).[3] CHARLS is a sister study of the US Health and Retirement Study (HRS) and the Survey of Health, Ageing and Retirement in Europe (SHARE). In addition to rich information on demographic and socio-economic characteristics, it provides longitudinal data on tested and self-reported health, which enables us to account for unobserved heterogeneities via panel data methods.

The sample was obtained via a multistage stratified probability proportional to size sampling design (Zhao et al., 2014). It comprises a national baseline survey conducted in 2011-2012 on 17,708 residents of 10,257 households in 450 villages and urban communities, with three follow-up interviews in 2013 (18,455 respondents), 2015 (20,967 respondents), and 2018 (20,813 respondents) (Zhao et al., 2021). In Section #4.4.6, we test whether the results are sensitive to attrition bias by using a 'variable addition test' (Verbeek, 2000; Verbeek & Nijman, 1992) and a three-year balanced panel.

Importantly, the data includes repeated measures on tested blood pressure, but also other biomarkers that enable the assessment of health perception. This will allow us to assess the robustness of our main findings by incorporating dyslipidemia and diabetes in addition to hypertension. In the 2011 and 2015 waves, venous blood samples were

[3] CHARLS is co-administered by the National School of Development and the Institute for Social Science Surveys at Peking University.

collected and analyzed (Chen et al., 2019) by first performing complete blood count (CBC) analyses at local county health centers and then sending the samples to the research headquarters for assay (Chen et al., 2019). Blood data were available for 11,847 respondents in 2011 and 13,420 respondents in 2015, respectively. The data further contains comprehensive information on in- and outpatient visits, which has been previously used in the literature to investigate healthcare utilisation (Chen & Ning, 2022; Tan et al., 2023; Zhou et al., 2020).

3.2 Sample construction

The study sample used for this study is adults aged 45 and above for whom systolic blood pressure (SBP), diastolic blood pressure (DBP), biomarkers (including total cholesterol (mg/dL), HDL cholesterol (mg/dL), and LDL cholesterol (mg/dL)) and healthcare utilisation information are available in the 2011-2015 waves. Our final samples range from 25,936 to 26,431 observations (see Table 3). It is worth highlighting that CHARLS includes community-dwelling individuals only, thus not providing information on nursing home residents. However, since less than 2% of Chinese older people live in nursing homes (Wu et al., 2018), this restriction is unlikely to affect our results.

3.3 Measuring health perception biases

Following the literature, we use the difference between objectively tested and subjectively reported health to assess health perception. Earlier work has explored health dimensions like mobility, cognition, high cholesterol, or hypertension (Arni et al., 2021; Spitzer & Shaikh, 2022; Spitzer et al., 2022; Spitzer & Weber, 2019). We focus on the latter, because hypertension is the leading risk for death in China (Z. W. Wang et al., 2018; Wu et al., 2023) and has not been explored in prior studies. In robustness analyses, we use alternative measures based on dyslipidemia and diabetes.

Objective hypertension is assessed based on physical examinations. DBP and SBP are tested three times and we use three measures to calculate the mean values of DBP and SBP. Then, following most of the existing literature, we define hypertension according to the International Diabetes Federation (IDF) cut-offs of DBP $\geq$ 90 mmHg or SBP $\geq$ 140 mmHg (Lei et al., 2012; Nie et al., 2023). Subjective hypertension is assessed based on self-reports, i.e. whether the respondent reported suffering from hypertension or not

(1=yes; 0=no).

We differentiate between individuals that overestimate their health (their subjective health is better than their objective health), underestimate their health (their subjective health is worse than their objective health), or achieve concordance (their subjective and objective health align). More specifically, respondents are considered to overestimate their health when they report no hypertension, but are hypertensive according to the physical examinations (Table 1). They are considered to underestimate their health when they subjectively report being hypertensive, but are not hypertensive according to the test. We further differentiate between positive and negative concordance. Respondents are categorized as positive concordant if they are not hypertensive according to the test and the self-report. They are categorized as negative concordant if they are hypertensive according to the test and the self-report. Further summary statistics are provided in Table 2.

<Insert Table 1 here>

The above definition of overestimating and underestimating health relates to the concept of overconfidence, which has been studied extensively in the psychology literature (Moore & Schatz, 2017; Murphy et al., 2017). It is commonly defined in three distinct ways: (i) overestimation of one's actual performance, (ii) overplacement of one's performance relative to others, and (iii) overprecision[4] in one's belief (Moore & Healy, 2008; Murphy et al., 2017). Classifications (i) and (iii) denote absolute overconfidence measures (Chen & Schildberg-Hörisch, 2019), while (ii) corresponds to relative overconfidence (Benoît & Dubra, 2011; Benoît et al., 2015). Since we assess health perception based on the difference between objectively tested and subjectively reported health, our measure of overconfidence adopts the first concept of overestimation. Moreover, our concept of health perception also allows us to consider underconfidence, i.e. the underestimation of health. Our measure thus differs from Nie et al. (2022) for China and Arni et al. (2021) for Germany, who mostly focus on relative health perception biases.[5] This strategy, which combines objectively tested and

---

[4] Overprecision indicates that the individual is overly certain of knowing the truth (Moore & Schatz, 2017).
[5] Relative health perception biases are defined as the difference between the subjectively perceived and the objectively measured rank in the population health distribution (Arni et al. 2021).

subjectively reported health measures, has been widely used in the existing literature on assessing health misperception (Arni et al., 2021; Nie et al., 2023; Spitzer et al., 2022; Spitzer & Weber, 2019).

3.4 Outcome variables: Healthcare utilization

Healthcare utilization is measured via six different outcome variables that cover a wide range of healthcare dimensions: the number of doctor visits in the past month, the number of hospital stays in the past year, out-of-pocket (OOP) outpatient cost[6], the total outpatient cost, OOP inpatient cost, and the total inpatient cost. Specifically, hospital stays denote the total number of visits to general hospitals, specialized hospitals, Chinese traditional medicine hospitals ("Zhongyi"), community healthcare centers, township hospitals, healthcare posts, village/private clinics and other healthcare organizations (Zhang et al., 2018). Total outpatient cost includes the total expenditure for outpatient care in the past month. OOP inpatient spending is the OOP expenditure for inpatient care in the past year, including fees paid to the hospital like ward fees, but excluding the wage of hired nurses, the fare, or rent (Zhang et al., 2018). Total inpatient cost is the total cost for inpatient care in the past year (Tan et al., 2018). All values are inflation-adjusted to 2015-standards.

Table 2 provides an overview of all outcome variables. On average, individuals in our sample visit a doctor once every two months (0.46 per month) and stay at the hospital once every five years (0.18). These values are similar to those reported by Zhang et al. (2018). The average monthly OOP outpatient cost, monthly total outpatient cost, yearly OOP inpatient cost, and yearly total inpatient cost are 147.6 yuan (23.7 US$), 239.0 yuan (38.4 US$), 701.8 yuan (112.67 US$) and 1203.9 yuan (193.3 US$), respectively.[7] For health misperceptions, approximately 12% of respondents underestimate their health whilst around 19% overestimate their health.

<Insert Table 2 here>

3.5 Control variables

---

[6] Based on the CHARLS question, outpatient care refers to visiting a public hospital, private hospital, public health center, clinic or health worker's or doctor's practice, or to home visits by a health worker or doctor.

[7] The 2015 average exchange rate of US dollars to RMB (yuan) is 6.2284.

Following Spitzer and Shaikh (2022) and Tan et al. (2018), we account for confounding variables by controlling for individual demographic and socio-economic characteristics, including age groups (45-54, 55-64 and 65+, with 45-54 as the reference group), gender (1 = male, 0 = female), education (measured on a 3-point scale: 1 = low: “at most primary school”, 2 = middle: “finished middle school, high school or vocational training”, and 3 = high: “two-/three-year college/associate degree, four-year college/bachelor degree, master or doctoral degree”), with “low-level education” as the reference group, [8] marital status (1 = married/partnered, 0 = other), number of limitations in activities of daily living (ADL), retirement (1 = yes, 0 = no) and household income. We control for retirement because it was shown to increase healthcare utilization in China (Zhang et al., 2018). As medical insurance is associated with increased healthcare utilization (Huang & Gan, 2017; Tan et al., 2018; Zhou et al., 2020), we also control for the health insurance scheme (UEBMI, URBMI, NRCMI, and Gongfei/PMI/Others), with no insurance[9] as the reference group.

Finally, given China’s diverse geography and its major urban-rural divide (Nie, Li, & Sousa-Poza, 2021; Zhang et al., 2018), we control for region (a categorical variable with 1 = East, 2 = Center, 3 = West, 4 = Northeast, with “East” as the reference group) and the current location of residence (1 = rural, 0 = urban).

Table A1 of the Appendix presents summary statistics for all control variables. Our sample consists of 48% men and 52% women. The low education group is the largest (89.2%), while 9.4% of the sample belong to the middle education group and only a small minority (1.4%) are highly educated. The vast majority (85.7%) of our sample are married or living together. 29.8% have retired, which is not surprising given the relatively high average age. Approximately 95% of the respondents in our sample are covered by medical insurance, which is similar to the values provided by Zhang et al. (2018). As regards the different health insurance schemes, the NRCMI (76.4%) is the most prevalent. Finally, the majority of respondents are rural residents (64.4%).

---

[8] A similar categorization has been used in Zhang et al. (2018) using the CHARLS dataset.

[9] No insurance means that individuals are not enrolled in any health insurance schemes. Since China's 2009 healthcare reform, insurance coverage has risen from 85% in 2008 to 96% in 2020. Despite improved access to healthcare, disparities remain. UEBMI offers the most generous coverage, while NRCMI provides limited outpatient care. Urban employees and retirees tend to have better resources compared to rural populations, highlighting ongoing inequalities in healthcare access (National Health Commission of the People's Republic of China, 2021).

3.6 Empirical strategy

3.6.1 Random effects Poisson estimation for doctor visits and hospital stays

The panel structure of CHARLS allows us to address unobserved heterogeneities between survey participants via panel data methods. While our approach does not entirely alleviate concerns regarding unobserved confounders and does not yield causal estimates, it enables us to account for unobserved differences between individuals and thus goes beyond previous evidence based on cross sectional data only (Arni et al., 2021; Spitzer & Shaikh, 2022).

To estimate how health perception biases are associated with doctor visits and hospital stays, we employ the following random effects (RE) Poisson model:

$$E(y_{it}|\alpha_i, x_{it}) = exp(\gamma_i + \theta HPB_{it} + x'_{it}\beta) = \alpha_i \, exp(\theta HPB_{it} + x'_{it}\beta), \gamma_i = ln\alpha_i \qquad (1)$$

where y_{it} denotes individual i's healthcare use (the number of doctor visits last month or the number of hospital stays last year) at time t, HPB_{it} represents individual i's health perception biases at time t, and x_{it} is a vector of demographic and socio-economic controls. We opt for a RE model over a fixed effects (FE) model for two primary reasons. First and foremost, there exists little variation in health perception within individuals across waves in our case. As emphasized by Seshamani and Gray (2004), FE estimators provide inconsistent estimates if there is not sufficient variation within individuals. Second, it is impossible for the FE model to provide estimated coefficients for time-invariant variables such as gender.

The above model is estimated separately for the subsample with objectively tested hypertension (negative concordance and overestimation) and the subsample without objectively tested hypertension (positive concordance and underestimation). As highlighted by Spitzer and Shaikh (2022), distinguishing those groups ensures that the reference group (concordance) has the same objective health status as those who misperceive their health (either overestimate or underestimate) – any confounding effects related to the health status are thus taken care of. As underestimation is only possible for respondents who are not objectively hypertensive, positive concordance is taken as the reference group for underestimation. In contrast, overestimation is only possible for those with tested hypertension. Here, negative concordance is taken as the reference category.

To explore further determinants of health perception, we also run a multinomial logit regression model, combining the subsamples with and without tested hypertension. The outcome variable is a 3-point categorical variable (1 = positive or negative concordance, 2 = underestimation and 3 = overestimation, with 1 as the reference category).

3.6.2 Random effects two-part model (RE-TPM) for healthcare expenditure

To estimate how health perception biases are associated with healthcare expenditure, we employ a random-effects two-part model (RE-TPM). TPMs have been growing in popularity and are considered state-of-the-art for dependent variables with severe skewness and substantial point mass at zero (Belotti et al., 2015).[10] Following the literature, we first estimate the probability that a respondent has any healthcare expenditures with a probit model using the full sample. In a second step, we then estimate a generalized linear model (GLM) based on respondents with healthcare expenditures larger than zero. In the second stage, a GLM with a gamma family and log link is usually applied to address the econometric problems caused by skewness in health care utilization studies (Manning & Mullahy, 2001). The log link is useful for correcting highly skewed data, while the gamma family may help to reduce heteroskedasticity concerns. We employ the following RE-TPM:

$$\Phi^{-1}[P(y_{it} > 0 | HPB_{it}, X_{it})] = \alpha_0 + \alpha_1 HPB_{it} + X'_{it}\theta + U_i \quad (2)$$

$$log[E(y_{it} | y_{it} > 0, HPB_{it}, X_{it})] = \beta_0 + \beta_1 HPB_{it} + X'_{it}\gamma + V_i \quad (3)$$

where equation 2 estimates the first-stage RE probit model and equation 3 estimates the second-stage RE GLM; y_{it} represents individual *i*'s healthcare spending (total and OOP expenditures for inpatient and outpatient care) at time *t*; HPB_{it} represents individual *i*'s health perception biases at time *t*; X_{it} is a vector of explanatory variables;

[10] Unlike instrumental variable (IV) regression or Heckman selection models, the TPM does not inherently require an exclusion restriction mainly because the two parts of the TPM are estimated separately, and the model does not explicitly account for the correlation between them. However, when addressing the concern about the endogeneity issue in either part of the model (e.g., unobserved confounding affecting both the decision to use healthcare and the level of healthcare utilization) using an IV approach or when modeling the correlation between the two parts using more advanced approaches such as a copula model, an exclusion restriction would be necessary. A detailed discussion is available in Belotti et al. (2015).

and U_i and V_i are random intercepts in the two equations for individual i and are assumed to be uncorrelated with X_{it}.

Following Nie and Li (2024), we employ the generalized structural equation modeling (GSEM) approach to perform RE-TPM estimation. It is worth noting that GSEM has three key features that are suitable for our estimation. First, equations in GSEM can take nonlinear forms, such as probit. Second, equations in GSEM can use different samples such as the full sample and subsamples, respectively. Third, individual-level RE can be specified as latent variables in GSEM. Like the Poisson model, the RE-TPM is estimated separately for the subsample with objectively tested hypertension and the subsample without objectively tested hypertension.

4. Results

4.1 Determinants of health perception biases

Tables A2 and A3 of the Appendix provide descriptive results separately for the subsample with objectively tested hypertension (negative concordance and overestimation) and the subsample without objectively tested hypertension (positive concordance and underestimation). Both health underestimation and overestimation increase with age and that men are more likely to overestimate their health than women (52% versus 45%, Tables A2 and A3).

Regarding determinants of health perception, results in Table A4 of the Appendix show that, relative to younger adults aged 45-54, older adults are more likely to misperceive their health both by underestimating (column 1) and overestimating (column 2). Men are less likely to underestimate their health (column 1) and more likely to be overestimate their health (column 2), which is in line with the findings by Spitzer and Weber (2019) for European adults. Findings for educational attainment, marital status, and retirement are inconclusive. Importantly, a higher number of ADLs is associated with a higher probability of health underestimation but a lower likelihood of overestimation, confirming our strategy to split the sample into those who are objectively healthy and those who are objectively unhealthy in the analyses.

4.2 Health misperception and healthcare utilization

Tables A2 of the Appendix show that average healthcare use and expenditures are significantly higher among individuals who underestimate their health compared to those with positive concordance. In contrast, individuals who overestimate their health have lower healthcare utilization and spending than those with negative concordance (Table A3). These findings are confirmed by our random effects Poisson estimation (Figure 1). For ease of interpretation, we report average marginal effects rather than estimated coefficients. Individuals who underestimate their hypertension visit the doctor more often and have more hospital stays than those who perceive their health correctly. Compared to respondents with positive concordance, those who underestimate their hypertension have approximately 0.30 more doctor visits and 0.30 more hospital stays. By contrast, respondents who overestimate their hypertension have only 0.43 and 0.55 of the predicted doctor visits and hospital stays, respectively, compared to those who perceive their health correctly.[11] These findings are in line with Spitzer and Shaikh (2022) for Europe, however, the effect sizes for doctor visit among Chinese middle-aged and older adults are much smaller compared to European individuals aged 50 and above, with around 25% (0.3/1.2) for underestimation and 24% (0.43/1.8) for overestimation, respectively. The results suggest that health underestimation is associated with more healthcare utilization, while health overestimation is related to less healthcare use, even after accounting for differences in health status.

<Insert Figure 1 here>

4.3 Health misperception and healthcare spending

We find that individuals who underestimate their hypertension have higher healthcare expenditure. This holds for OOP and total expenditures, both for outpatient and inpatient care (Figure 2a). Specifically, individuals with hypertension underestimation have approximately 49 yuan/month (8 US$/month) higher OOP outpatient spending, 81 yuan/month (13 US$/month) higher total outpatient spending, 222 yuan/year (36 US$/year) higher OOP inpatient spending, and 428 yuan/year (69 US$/year) higher total inpatient spending. By contrast, those who overestimate their hypertension are less likely to have at least some expenditure (Figure 2b). Specifically, individuals with hypertension overestimation have approximately 61 yuan/month (10 US$/month) lower

[11] The estimates for the control variables can be found in the Appendix Table A5.

OOP outpatient spending, 98 yuan/month (16 US$/month) lower total outpatient spending, 294 yuan/year (47 US$/year) lower OOP inpatient spending, and 498 yuan/year (80 US$/year) lower total inpatient spending.[12] Additionally, as a robustness check, we also rerun the regression by employing a GLM with a gaussian family and identity link in the second stage of RE-TPM estimation. Once again, results from Appendix Table A7 and Figure A1 confirm that our main findings are quantitively similar and robust.

<Insert Figure 2 here>

4.4 The mediating role of children and insurances

To assess whether the relationship between health misperceptions and healthcare utilization is moderated by the number of co-resident children, we conduct three additional analyses: (i) incorporating interaction terms between health underestimation/overestimation and the number of co-resident children, (ii) examining interaction terms with the number of co-resident sons, and (iii) analyzing interaction terms with the number of co-resident daughters. Given the urban-rural disparity, we also re-estimate the models separately for rural and urban areas.

The results indicate that individuals who underestimate their hypertension are more likely to visit a doctor. However, in urban areas, the interaction between health underestimation and the number of co-resident children—particularly sons—is negative (Column 2, Panels A and B, Table A8). This suggests that the presence of co-resident sons in urban China weakens the positive effect of health underestimation on doctor visits. These findings provide evidence for the moderating role of co-resident sons in the relationship between health misperception and healthcare utilization in China.

The substantial variations in benefit package, reimbursement rate, and cost-sharing

---

[12] The results for all the controls are available Appendix Table A6.

across China's health insurance schemes - UEMBI, URBMI, and NRCMI - may result in heterogeneous healthcare expenditures. To account for this, we rerun the RE-TPM estimates for each scheme (see Table A9). Three key findings are worth mentioning. First, overall, individuals who underestimate their hypertension exhibit higher healthcare expenditure, whereas those who overestimate their health have lower expenditure. However, most estimates for UEMBI (Figures 3a and 3b) and URBMI (Figures 3c and 3d) are statistically insignificant due possibly to limited sample sizes. Second, the marginal effects of health underestimation on outpatient expenditure are significantly larger for NRCMI enrollees compared to URBMI participants. This discrepancy aligns with institutional differences: URBMI does not cover outpatient services but NRCMI partially covers them (Nie and Zhao, 2023). Consequently, NRCMI's outpatient benefits amplify the expenditure response to health underestimation compared to URBMI. Third, the effect size of health overestimation on OOP expenditures is markedly stronger for NRCMI beneficiaries than for UEMBI and URBMI enrollees. A plausible explanation lies in the higher annual reimbursement caps under NRCMI, which may reduce upfront financial barriers but increase OOP costs for those who delay care due to overoptimistic health assessments.[13]

<Insert Figure 3 here>

---

[13] We also conducted a number of other analyzes on subsamples, including by gender, age groups, poverty status (poor vs. nonpoor), and region (urban vs. rural area) (see Appendix Table A10-A13). Several findings are worth mentioning: (1) For gender, the effect of health misperception on healthcare utilization is stronger for men than for women, but the differences are not statistically significant (see Table A10). The only exception is the impact of health overestimation on hospital stays ($p<0.01$), where we find statistically significant stronger effects of men versus women. This finding may be related to the pronounced gender difference in healthcare utilization, as men use less healthcare services than women (Wang et al., 2018). (2) For age groups, younger adults aged 45-54 years who overestimate their hypertension use more healthcare services, in particular hospital stays, compared to older adults (see Table A11). However, these results are statistically insignificant, except for doctor visits among those aged 55-64 vs. aged 65+ ($p<0.05$). (3) For poverty status, a stronger impact of health overestimation on healthcare utilization is observed among the poor compared to the non-poor, but a Wald test shows that the differences are statistically insignificant (Table A12). (4) For regions, we consistently observe stronger effects on hospital stays among those who underestimate their hypertension in urban area compared to rural area while the effects of health overestimation are more pronounced in rural areas (Table A13). These differences are statistically significant ($p<0.01$).

4.4 Robustness checks

Our findings remain robust across various sensitivity analyses. Specifically, we address endogeneity concerns, including reverse causality, account for unobserved individual heterogeneity and overdispersion, employ alternative measures of health perception, mitigate sample attrition, and try to address the role that medical check-ups may influence perceptions.

4.4.1 Endogeneity issues

While we are not aiming at causal estimates, endogeneity, particularly reverse causality, remains a concern in our analysis. Although health perception influences healthcare utilization, the reverse may also hold, as frequent doctor visits could enhance health concordance (Spitzer and Shaikh, 2022). To mitigate this issue, we investigate the relationship between health misperception in 2011 and subsequent healthcare utilization in 2015. The results, presented in Appendix Table A14, largely support our main findings, except for health underestimation in relation to doctor visits, which remains positive but statistically insignificant.

4.4.2 Alternative estimation methods

We provide within-estimations using individual fixed effects to account for potential unobserved individual heterogeneity that may affect both health perception and healthcare utilization. Although the sample sizes are significantly smaller compared to the RE Poisson estimation due to the little within-individual variation in health perceptions across waves, the results from fixed effects Poisson and fixed effects negative binomial estimates confirm the robustness of our key findings, except for the association between health underestimation and hospital stays (see Table A15, Panels A and B).To deal with the overdispersion[14], following Liu et al. (2021), we also introduce a RE negative binomial regression model to re-estimate the linkage between health perception biases and healthcare utilization. The new results confirm that health underestimation is linked to more doctor visits and hospital stays compared to positive concordance, and that health overestimation is associated with less doctor visits and hospitals stays than negative concordance (see Table A15, Panel C). The results are

[14] An unusual property of the Poisson distribution is that the mean and variance are equal, thereby resulting in a problem of overdispersion.

quantitatively similar to the main estimates in Table 4 using RE Poisson estimation.

4.4.3 Check-ups and perceptions

Individuals who undergo regular check-ups may have more accurate health information, potentially reducing biases in health perception. Although direct measures of recent or frequent check-ups are unavailable, we conduct two additional robustness checks to address this concern: (i) examining whether the type of the last outpatient visit (1= a follow-up visit, and 0= the first visit) is linked to health misperception, and (ii) assessing whether the purpose of the last outpatient visit (categorized as 1 = medical check-up, and 0 = others) influences health misperception. In the former case, one could assume that a follow-up visit may proxy the frequency of a check-up, whereas in the latter case we can identify whether the last visit was indeed a check-up. Of course, both measures have drawbacks in that they cannot precisely measure the frequency of check-up, nor when they last occurred. Results indicate that a follow-up outpatient visit, compared to a first visit, is insignificantly associated with health underestimation but significantly and positively related to health overestimation (Panel A, Table A16), suggesting that individuals who undergo regular check-ups may reduce their health misperception especially overestimation. Regarding the purpose of the last outpatient visit, results show no significant association with health misperception, whether underestimation or overestimation (Panel B, Table A16).

4.4.4 Unobserved heterogeneity

Since the CHARLS survey does not allow respondents to report confidence intervals, they can only provide point estimates. This implies that every individual has a certain probability of experiencing concordance, overestimation, or underestimation due to estimation error. As a result, unobserved heterogeneity in health self-assessment remains a concern. To address this issue, we employ two approaches: First, we use a latent class Poisson model to classify individuals into latent groups based on the underlying distribution of their health status estimates. Each group is characterized by distinct means and variances, capturing differences in the precision and bias of their health assessments. This approach allows us to model the probability of health concordance, underestimation, or overestimation as a function of latent group membership. The results confirm the robustness of our main findings (see Panel A,

Table A17). Second, we apply a random-coefficient model (multilevel mixed-effects Poisson regression) to estimate individual-level variance in health misperception effects. By incorporating a random slope for health misperception, we account for individual-specific deviations in misperception bias (mean) and precision (variance). Again, the results (see Panel B, Table A17) reinforce the robustness of our main findings.[15]

4.4.5 Alternative measures of health misperception

As an additional robustness check, we measure health perception biases using dyslipidemia and diabetes, thereby following Nie et al. (2023). Specifically, we define dyslipidemia as total cholesterol ≥ 240 mg/dL or HDL cholesterol < 40 mg/dL or LDL cholesterol ≥ 160 mg/dL or triglycerides > 200 mg/dL (Pan et al., 2016), and diabetes as HbA1c ≥ 6.5 % or taking diabetes medication (Chen et al., 2019). RE Poisson results indicate that individuals who underestimate their dyslipidemia or diabetes are more likely to use healthcare services, while those who overestimate their dyslipidemia or diabetes are less likely (see Table A19). Hence, our results are robust to the use of different health perception measures.[16]

4.4.6 Attrition bias

As the CHARLS dataset is an unbalanced panels with respondents being included for different lengths of time, we are addressing potential attrition biases in a separate robustness check (Verbeek, 2000). In particular, we address the possibility that some respondents left the sample for reasons correlated with health or healthcare utilization (Zhang et al., 2018). For this, we employ a "variable addition test" (Verbeek, 2000; Verbeek & Nijman, 1992), which is evaluated on the significance of an added variable (defined as the number of surveyed years that each respondent is present in the survey),

---

[15] Accounting for heteroskedasticity is critical, particularly in non-linear limited-dependent models like our random-effects Poisson estimation, as it can lead to biased estimates. To address this concern, we have performed two additional analyses: (1) We have employed cluster-robust standard errors in our random-effects Poisson estimation to account for potential heteroskedasticity across individuals and over time. This adjustment ensures that our inference remains valid even if the variance differs across groups (e.g., individuals with health concordance, underestimation, or overestimation) or over different waves of our panel. The results confirm the robustness of our main findings (see Panel A, Table A18). (2) We have employed generalized linear models with heteroskedasticity-consistent variance estimators to account for potential heteroskedasticity. The results, once again, show that our main findings are robust (see Panel B, Table A18).

[16] We also replace ADLs with self-assessed health (SAH) as the proxy of individual health mainly because SAH encompasses not only mental and physical health but also subjective experience of acute and chronic diseases and overall feelings of well-being. We then rerun the estimates and show that our main findings remain robust (see Table A20).

and its insignificance indicating the nonexistence of attrition bias. The results from this test are insignificant for both doctor visits and hospital stays (Panel A, Table A21), leading to the conclusion that endogenous attribution bias is unlikely an issue. Additionally, we also employ a three-year balanced panel and rerun the estimates. Reassuringly, the new results are consistent with main findings (Panel B, Table A21).

5. Discussion and conclusions

This study has investigated the role of health perception biases in shaping healthcare utilization and related expenditures in China—an underexplored yet crucial context characterized by rapid population aging, fragmented healthcare coverage, and strong family-based care traditions. Although growing research has highlighted how biased health perceptions affect individual health behaviors, evidence on how these biases translate into healthcare utilization and costs, particularly in developing countries, has remained limited. Our study addresses this gap by analyzing nationally representative longitudinal data from the 2011–2015 China Health and Retirement Longitudinal Study (CHARLS), enabling us to account for unobserved individual heterogeneity and explore key social and institutional mechanisms that shape these relationships.

The study yields several key findings that are robust to a wide range of sensitivity analyses. First, we show that approximately 44% of the Chinese aged 45 and older overestimate their health and around 18% of them underestimate their health. These values are much higher than those reported by Spitzer and Weber (2019), which is surprising given their older sample. One explanation could be that our health perception measure is based on hypertension, while Spitzer and Weber use mobility and cognition. A second rational could be that uneven healthcare access in China results in lower levels of diagnosed hypertension among older adults (Lei et al., 2012), which could explain the very high levels of health overestimation in China. Similar to earlier work on Europe (Spitzer & Shaikh, 2022; Spitzer & Weber, 2019), we also find that health misperception increases with age.

Second, our results show that individuals who underestimate their hypertension visit the doctor more often and have more hospital stays, even after accounting for differences in health status. By contrast, those who overestimate their health are less likely to use those healthcare services. This observation is consistent with the findings by Spitzer and Shaikh (2022) for Europe. However, the effect sizes for doctor visit among the Chinese middle-aged and older adults are much smaller compared to European individuals aged 50 and above, with around 25% (0.3/1.2) for underestimation and 24% (0.43/1.8) for overestimation, respectively. The inclination to visit a doctor in China may be lower than in Europe due to variations in healthcare systems, cultural attitudes, and financial considerations. China's hospital-centric system means that many patients bypass primary care and go directly to hospitals, often leading to overcrowding and long wait times, which can discourage routine visits. Cultural factors also play a role—while Europeans are generally more proactive about regular check-ups, many Chinese individuals, particularly in rural areas, may delay seeking medical attention due to cost concerns, reliance on traditional medicine, or the perception that medical care should be sought only when symptoms become severe (Lu et al., 2017). Despite near-universal insurance coverage in China, high out-of-pocket expenses can also deter people from seeking care.

Our results also show that individuals who underestimate their hypertension are more likely to see the doctor, and that its interaction term with the number of co-resident adult children, especially sons, in urban area is negative, suggesting that the existence of co-resident sons in urban China would attenuate the positive impact of health underestimation on visiting a doctor. This provides the evidence of the moderating effect of co-resident sons on the linkage between health misperception and healthcare utilization in China. This underlines the important role that the family plays in healthcare decisions, with children often influencing their parents' willingness to seek treatment (Bowman & Singer, 2001). This family-centered approach is deeply rooted in Confucian familism, which emphasizes collective decision-making and can influence the patient's willingness to seek treatment (Yang et al., 2022).

Finally, we provide novel results on the linkage between health perception biases and healthcare expenditures (OOP and total costs for outpatient and inpatient services). Individuals who underestimate their hypertension are more likely to have higher healthcare expenditure, with approximately 49 yuan/month (8 US$/month) for OOP outpatient spending, 81 yuan/month (13 US$/month) for total outpatient spending, 222 yuan/year (36 US$/year) for OOP inpatient spending, and 428 yuan/year (69 US$/year) for total inpatient spending). Additionally, those who overestimate their hypertension are less likely to have at least some expenditure, with around 61 yuan/month (10 US$/month) for OOP outpatient spending, 98 yuan/month (16 US$/month) for total outpatient spending, 294 yuan/year (47 US$/year) for OOP inpatient spending, and 498 yuan/year (80 US$/year) for total inpatient spending. An interesting finding of our analysis is that the extent to which health misperceptions affect healthcare expenditures is largely influenced by the health insurance scheme. In particular, health underestimation has a larger impact on outpatient spending for NRCMI enrollees than for URBMI participants, as NRCMI's partial outpatient coverage amplifies expenditure responses. Additionally, health overestimation leads to higher out-of-pocket costs for NRCMI beneficiaries compared to UEMBI and URBMI enrollees, likely due to NRCMI's higher reimbursement caps reducing upfront barriers but increasing costs for delayed care.

Our study has at least three key limitations. First, while we have controlled for a comprehensive set of sociodemographic characteristics and employed several rigorous empirical strategies to address potential endogeneity in health perception, we cannot fully establish a causal relationship between health misperceptions and healthcare utilization. However, our analysis is the first to leverage longitudinal panel data, enabling us to mitigate several biases that previous cross-sectional studies could not address. Second, health literacy—encompassing knowledge, behaviors, and skills—is a crucial determinant of health outcomes and healthcare-related behaviors (Paasche-Orlow & Wolf, 2007). Individuals with low health literacy often struggle to understand medical terminology, assess health risks, and make informed healthcare decisions, which can lead to underuse or non-use of medical services (Xu et al., 2022). Unfortunately, due to data limitations, we are unable to directly control for health literacy in our empirical analysis.[17] Third, individuals display varying risk attitudes, ranging from pessimistic and risk-averse to optimistic and risk-tolerant, which may influence healthcare utilization patterns. However, due to data constraints, we are unable to separate the effects of risk attitudes from the broader unobserved

[17] Since CHARLS does not provide detailed information on health knowledge and health literacy, we utilize data from the 2004, 2006, and 2009 waves of the China Health and Nutrition Survey (CHNS) to examine whether knowledge of dietary guidelines is associated with health misperception among middle-aged and older adults. We define health misperception using the same methodology as in this study. The results indicate that knowledge of dietary guidelines is not significantly associated with health underestimation but is significantly and negatively correlated with health overestimation, albeit at a 10% significance level. This suggests that health literacy may play a role in reducing the overestimation of health. Full results are available upon request.

heterogeneity in health misperceptions. Notably, a recent study from Germany (Arni et al., 2021) finds that risk tolerance—whether risk-averse or risk-loving—is not significantly associated with health misperceptions, regardless of whether individuals overestimate or underestimate their health. This suggests that risk attitudes may have a limited impact on shaping health misperceptions. Further research is needed to examine this relationship more comprehensively, utilizing richer datasets and experimental designs to better isolate the underlying causal mechanisms.

These findings have important policy implications. First, targeted health literacy campaigns and the promotion of regular health check-ups and screenings may be important to increase the healthcare utilization of those who overestimate their health. Second, with the increasing process of population aging, older adults are facing numerous challenges in healthcare utilizations and services. Given the important role of co-resident children in mitigating the impact of health underestimation on healthcare expenditure, recognizing the pivotal role of intergenerational support from adult children in reducing health misperception and avoiding unnecessary healthcare spending of older adults is essential for establishing an age-friendly and sustainable healthcare system.

Finally, while such health literacy campaigns may also be important for tackling the higher healthcare utilization of those who underestimate their health, additional efforts may be needed to provide effective health counseling and regulatory measures to avoid unnecessary demand-side driven healthcare use. Moreover, such measures may be an important lever for curbing the waste of limited medical resources due to over-testing and over-treatment, which is prevalent in China's fee-for-service payment system and a major challenge for the efficient provision of healthcare services. Our results also show the importance that health insurance schemes play in mediating the effects of health misperceptions on healthcare expenditures. These findings underscore the critical role of China's ongoing health insurance reforms—such as integrating urban-rural schemes and refining reimbursement structures—in addressing disparities linked to health misperceptions. Targeted policy measures, including expanding outpatient coverage under URBMI and calibrating NRCMI's reimbursement caps to balance access and financial risk, could advance the *Healthy China 2030* goal of universal, equitable, and financially sustainable healthcare.

References

Akindele, M. O., Phillips, J., Igumbor, E., & Useh, U. (2017). Body weight misperception and dissatisfaction among overweight and obese adult Nigerians. *JMIR Public Health and Surveillance, 3*(3), e56. https://doi.org/10.2196/publichealth.7047

Arni, P., Dragone, D., Goette, L., & Ziebarth, N. R. (2021). Biased health perceptions and risky health behaviors-Theory and evidence. *Journal of Health Economics, 76*, 102425. https://doi.org/https://doi.org/10.1016/j.jhealeco.2021.102425

Belotti, F., Deb, P., Manning, W. G., & Norton, E. C. (2015). Twopm: two-part models. *The Stata Journal, 15*(1), 3-20. https://doi.org/10.1177/1536867X1501500102

Benoît, J.-P., & Dubra, J. (2011). Apparent overconfidence. *Econometrica, 79*(5), 1591-1625. https://doi.org/10.3982/ECTA8583

Benoît, J.-P., Dubra, J., & Moore, D. A. (2015). Does the better-than-average effect show that people are overconfident? Two experiments. *Journal of the European Economic Association, 13*(2), 293-329. https://doi.org/10.1111/jeea.12116

Bowman, K. W., & Singer, P. A. (2001). Chinese seniors' perspectives on end-of-life decisions. *Social Science & Medicine, 53*(4), 455-464. https://doi.org/https://doi.org/10.1016/S0277-9536(00)00348-8

Caliendo, F., & Huang, K. X. D. (2008). Overconfidence and consumption over the life cycle. *Journal of Macroeconomics, 30*(4), 1347-1369. https://doi.org/https://doi.org/10.1016/j.jmacro.2007.10.001

Chen, H., & Ning, J. (2022). The impacts of long-term care insurance on health care utilization and expenditure: evidence from China. *Health Policy and Planning*, czac003. https://doi.org/10.1093/heapol/czac003

Chen, S., & Schildberg-Hörisch, H. (2019). Looking at the bright side: the motivational value of confidence. *European Economic Review, 120*, 103302. https://doi.org/https://doi.org/10.1016/j.euroecorev.2019.103302

Chen, X., et al. (2019). Venous blood-based biomarkers in the China Health and Retirement Longitudinal Study: Rationale, design, and results from the 2015 wave. *American Journal of Epidemiology, 188*(11), 1871-1877. https://doi.org/10.1093/aje/kwz170

Dieleman, J. L., et al. (2017). Future and potential spending on health 2015–40: development assistance for health, and government, prepaid private, and out-of-pocket health spending in 184 countries. *The Lancet, 389*(10083), 2005-2030. https://doi.org/https://doi.org/10.1016/S0140-6736(17)30873-5

Friehe, T., & Pannenberg, M. (2019). Overconfidence over the lifespan: Evidence from Germany. *Journal of Economic Psychology, 74*, 102207. https://doi.org/https://doi.org/10.1016/j.joep.2019.102207

Heard, C., Scuffham, P. A., Ratcliffe, J., & Whitty, J. A. (2017). The association between misperceptions around weight status and quality of life in adults in Australia. *Health and Quality of Life Outcomes, 15*(1), 53.

https://doi.org/10.1186/s12955-017-0627-7
Hou, S., Niu, Y., & Yang, J. (2018). Optimal consumption-portfolio rules with biased beliefs. *Economics Letters, 173*, 152-157. https://doi.org/https://doi.org/10.1016/j.econlet.2018.10.003
Hu, B. (2019). Projecting future demand for informal care among older people in China: the road towards a sustainable long-term care system. *Health Economics, Policy and Law, 14*(1), 61-81. https://doi.org/10.1017/S1744133118000221
Huang, F., & Gan, L. (2017). The impacts of China's urban employee basic medical insurance on healthcare expenditures and health outcomes. *Health Economics, 26*(2), 149-163. https://doi.org/https://doi.org/10.1002/hec.3281
Lei, X., Song, G. X., & Su, X. J. (2021). *Information, belief, and health behavior: evidence from China. Department of Economics, University of Alberta Working Paper No. 2021-09*. Department of Economics, University of Alberta.
Lei, X., Yin, N., & Zhao, Y. (2012). Socioeconomic status and chronic diseases: the case of hypertension in China. *China Economic Review, 23*(1), 105-121. https://doi.org/https://doi.org/10.1016/j.chieco.2011.08.004
Lim, H., & Wang, Y. (2013). Body weight misperception patterns and their association with health-related factors among adolescents in South Korea. *Obesity, 21*(12), 2596-2603. https://doi.org/https://doi.org/10.1002/oby.20361
Liu, D., et al. (2021). Investigating the relationship between social care supply and healthcare utilization by older people in England. *Health Economics, 30*(1), 36-54. https://doi.org/https://doi.org/10.1002/hec.4175
Lu, H., et al. (2017). Healthcare seeking behaviour among Chinese elderly. *International Journal of Health Care Quality Assurance, 30*(3), 248-259. https://doi.org/10.1108/IJHCQA-10-2015-0132
Malmendier, U., & Tate, G. (2015). Behavioral CEOs: the role of managerial overconfidence. *Journal of Economic Perspectives, 29*(4), 37-60. https://doi.org/10.1257/jep.29.4.37
Manning, W. G., & Mullahy, J. (2001). Estimating log models: to transform or not to transform? *Journal of Health Economics, 20*(4), 461-494. https://doi.org/https://doi.org/10.1016/S0167-6296(01)00086-8
Mazzonna, F., & Peracchi, F. (2024). Are older people aware of their cognitive decline? Misperception and financial decision-making. *Journal of Political Economy, 132*(6), 1793-1830. https://doi.org/10.1086/728697
McGraw, A. P., Mellers, B. A., & Ritov, I. (2004). The affective costs of overconfidence. *Journal of Behavioral Decision Making, 17*(4), 281-295. https://doi.org/https://doi.org/10.1002/bdm.472
Moore, D. A., & Healy, P. J. (2008). The trouble iwth overconfidence. *Psychological Review, 115*(2), 502-517.
Moore, D. A., & Schatz, D. (2017). The three faces of overconfidence. *Social and Personality Psychology Compass, 11*(8), e12331. https://doi.org/10.1111/spc3.12331
Murphy, S. C., Barlow, F. K., & von Hippel, W. (2017). A longitudinal test of three

theories of overconfidence. *Social Psychological and Personality Science, 9*(3), 353-363. https://doi.org/10.1177/1948550617699252

National Health Commission of the People's Republic of China. (2021). *China Health Statistics Yearbook 2020*.

Nie, P., & Li, Q. (2024). Does energy poverty increase health care expenditures in China? *Applied Economics, 56*(35), 4209-4235. https://doi.org/10.1080/00036846.2023.2210823

Nie, P., Li, Q., & Sousa-Poza, A. (2021). Energy poverty and subjective well-being in China: new evidence from the China Family Panel Studies. *Energy Economics, 103*, 105548. https://doi.org/https://doi.org/10.1016/j.eneco.2021.105548

Nie, P., et al. (2021). The change and correlates of healthy ageing among Chinese older adults: findings from the China health and retirement longitudinal study. *BMC Geriatrics, 21*(1), 78. https://doi.org/10.1186/s12877-021-02026-y

Nie, P., et al. (2022). "The better you feel, the harder you fall": health perception biases and mental health among Chinese adults during the COVID-19 pandemic. *China Economic Review, 71*, 101708. https://doi.org/https://doi.org/10.1016/j.chieco.2021.101708

Nie, P., Wang, L., & Sousa-Poza, A. (2023). Health perception biases and risky health behaviours in China. *Applied Economics Letters, 30*(2), 221-228. https://doi.org/10.1080/13504851.2021.1983125

Nie, P., & Zhao, Y. H. (2023). Aging in China. In D. E. Bloom, A. Sousa-Poza, & U. Sunde (Eds.), *Routledge Handbook on the Economics of Ageing* (pp. 691-712). Routledge.

Ortoleva, P., & Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review, 105*(2), 504-535. https://doi.org/10.1257/aer.20130921

Paasche-Orlow, M. K., & Wolf, M. S. (2007). The causal pathways linking health literacy to health outcomes. *Am J Health Behav, 31 Suppl 1*, S19-26. https://doi.org/10.5555/ajhb.2007.31.supp.S19

Park, S., Lee, S., Hwang, J., & Kwon, J.-W. (2017). The impact of weight misperception on health-related quality of life in Korean adults (KNHANES 2007–2014): a community-based cross-sectional study. *BMJ Open, 7*(6), e016098. https://doi.org/10.1136/bmjopen-2017-016098

Pirinsky, C. (2013). Confidence and economic attitudes. *Journal of Economic Behavior & Organization, 91*, 139-158. https://doi.org/https://doi.org/10.1016/j.jebo.2013.04.013

Saleem, M. D., Ahmed, G., Mulla, J., Haider, S. S., & Abbas, M. (2013). Weight misperception amongst youth of a developing country: Pakistan -a cross-sectional study. *BMC Public Health, 13*(1), 707. https://doi.org/10.1186/1471-2458-13-707

Seshamani, M., & Gray, A. M. (2004). A longitudinal study of the effects of age and time to death on hospital costs. *Journal of Health Economics, 23*(2), 217-235. https://doi.org/https://doi.org/10.1016/j.jhealeco.2003.08.004

Silva, S. U. d., et al. (2021). Association between body weight misperception and dietary patterns in Brazilian adolescents: cross-sectional study using ERICA data. *PLOS ONE, 16*(9), e0257603. https://doi.org/10.1371/journal.pone.0257603

Spitzer, S., & Shaikh, M. (2022). Health misperception and healthcare utilisation among older Europeans. *The Journal of the Economics of Ageing, 22*, 100383. https://doi.org/https://doi.org/10.1016/j.jeoa.2022.100383

Spitzer, S., Shaikh, M., & Weber, D. (2022). Older Europeans' health perception and their adaptive behaviour during the COVID-19 pandemic. *European Journal of Public Health, 32*(2), 322-327. https://doi.org/10.1093/eurpub/ckab221

Spitzer, S., & Weber, D. (2019). Reporting biases in self-assessed physical and cognitive health status of older Europeans. *PLOS ONE, 14*(10), e0223526. https://doi.org/10.1371/journal.pone.0223526

Tan, S., et al. (2023). Trends in blood pressure and hypertension among older adults and oldest-old individuals in China between 2008-2018. *Hypertension Research, 46*(5), 1145-1156. https://doi.org/10.1038/s41440-023-01183-4

Tan, S. Y., Wu, X., & Yang, W. (2018). Impacts of the type of social health insurance on health service utilisation and expenditures: implications for a unified system in China. *Health Economics, Policy and Law, 14*(4), 468-486. https://doi.org/10.1017/S174413311800018X

Tiefenbeck, V., et al. (2016). Overcoming salience bias: how real-time feedback fosters resource conservation. *Management Science, 64*(3), 1458-1476. https://doi.org/10.1287/mnsc.2016.2646

Verbeek, M. (2000). *A Guide to Modern Econometrics*. Wiley.

Verbeek, M., & Nijman, T. (1992). Testing for selectivity bias in panel data models. *International Economic Review, 33*(3), 681-703. https://doi.org/10.2307/2527133

Wang, S., Marquez, P., & Langenbunner, J. (2011). *Toward a healthy and harmonious life in China: stemming the rising tide of non-communicable diseases*.

Wang, Z., Li, X., Chen, M., & Si, L. (2018). Social health insurance, healthcare utilization, and costs in middle-aged and elderly community-dwelling adults in China. *International Journal for Equity in Health, 17*(1), 17. https://doi.org/10.1186/s12939-018-0733-0

Wang, Z. W., et al. (2018). Status of hypertension in China. *Circulation, 137*(22), 2344-2356. https://doi.org/10.1161/CIRCULATIONAHA.117.032380

Wen, C., & Zhang, Q. (2023). The role of intergenerational support in shaping oral healthcare-seeking behavior among older adults in China. *Frontiers in Public Health, 11*. https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2023.1234539

Wu, C., Newman, A. B., Dong, B.-R., & Odden, M. C. (2018). Index of Healthy Aging in Chinese older adults: China Health and Retirement Longitudinal Study. *Journal of the American Geriatrics Society, 66*(7), 1303-1310. https://doi.org/10.1111/jgs.15390

Wu, N., Xie, X., Cai, M., Han, Y., & Wu, S. (2023). Trends in health service needs, utilization, and non-communicable chronic diseases burden of older adults in

China: evidence from the 1993 to 2018 National Health Service Survey. *International Journal for Equity in Health, 22*(1), 169. https://doi.org/10.1186/s12939-023-01983-7

Xu, L.-m., Xie, L.-f., Li, X., Wang, L., & Gao, Y.-m. (2022). A meta-analysis of factors influencing health literacy among Chinese older adults. *Journal of Public Health, 30*(8), 1889-1900. https://doi.org/10.1007/s10389-021-01638-3

Yang, Y., Qu, T., Yang, J., Ma, B., & Leng, A. (2022). Confucian familism and shared decision making in end-of-life care for patients with advanced cancers. *International Journal of Environmental Research and Public Health, 19*(16).

You, Y., Wang, J., Teng, W., Ma, G., & Liu, P. (2018). Blood pressure and noncommunicable diseases in middle-aged and older adults in China. *PLOS ONE, 13*(11), e0206635. https://doi.org/10.1371/journal.pone.0206635

Yu, Y., et al. (2022). Association between living arrangements and cognitive decline in older adults: a nationally representative longitudinal study in China. *BMC Geriatrics, 22*(1), 843. https://doi.org/10.1186/s12877-022-03473-x

Zhang, A., Nikoloski, Z., & Mossialos, E. (2017). Does health insurance reduce out-of-pocket expenditure? Heterogeneity among China's middle-aged and elderly. *Social Science & Medicine, 190*, 11-19. https://doi.org/https://doi.org/10.1016/j.socscimed.2017.08.005

Zhang, Y., Salm, M., & van Soest, A. (2018). The effect of retirement on healthcare utilization: evidence from China. *Journal of Health Economics, 62*, 165-177. https://doi.org/https://doi.org/10.1016/j.jhealeco.2018.09.009

Zhao, Y., Smith, J. P., & Strauss, J. (2014). Can China age healthily? *The Lancet, 384*(9945), 723-724. https://doi.org/https://doi.org/10.1016/S0140-6736(14)61292-7

Zhao, Y. H., et al. (2021). *China Health and Retirement Report*.

Zhou, B., et al. (2019). Long-term and recent trends in hypertension awareness, treatment, and control in 12 high-income countries: an analysis of 123 nationally representative surveys. *The Lancet, 394*(10199), 639-651. https://doi.org/10.1016/S0140-6736(19)31145-6

Zhou, Y., et al. (2020). Medical insurance and healthcare utilization among the middle-aged and elderly in China: evidence from the China health and retirement longitudinal study 2011, 2013 and 2015. *BMC Health Services Research, 20*(1), 654. https://doi.org/10.1186/s12913-020-05522-w

Tables

Table 1 Health perception categories using hypertension

	Objective hypertension (clinical test)	
Subjective hypertension (self-report)	Non-hypertensive	Hypertensive
Non-hypertensive	Positive concordance	Overestimation
Hypertensive	Underestimation	Negative concordance

Note: Objective hypertension is defined according to the International Diabetes Federation (IDF) cut points (diastolic blood pressure ≥ 90 mmHg or systolic blood pressure ≥ 140 mmHg). Subjective hypertension is defined based on whether the respondent reported as having hypertension or not. Positive concordance includes individuals who are objectively hypertensive and also report being hypertensive. Negative concordance involves individuals who are neither objectively nor subjectively hypertensive.

Table 2 Summary statistics for outcome variables and health misperceptions: CHARLS 2011-2015

	N	Mean	S.D.	Min.	Max.	Median
Outcome variables						
Number of doctor visits	26,042	0.463	1.454	0	30	0
Number of hospital stays	26,431	0.175	0.615	0	20	0
OOP cost (outpatient)	25,936	147.558	1426.474	0	103000	0
Total cost (outpatient)	26,047	238.961	5261.912	0	800000	0
OOP cost (inpatient)	26,244	701.760	4765.259	0	174000	0
Total cost (inpatient)	26,213	1203.884	7082.466	0	214000	0
Health misperceptions						
Positive concordance	26,431	0.549	0.498	0	1	1
Underestimation	26,431	0.118	0.323	0	1	0
Overestimation	26,431	0.146	0.353	0	1	0
Negative concordance	26431	0.186	0.389	0	1	0
Objective and subjective hypertension						
Objective hypertension	26431	0.332	0.471	0	1	0
Subjective hypertension	26431	0.304	0.460	0	1	0

Note: OOP=out-of-pocket. Sampling weights are applied.

Figures

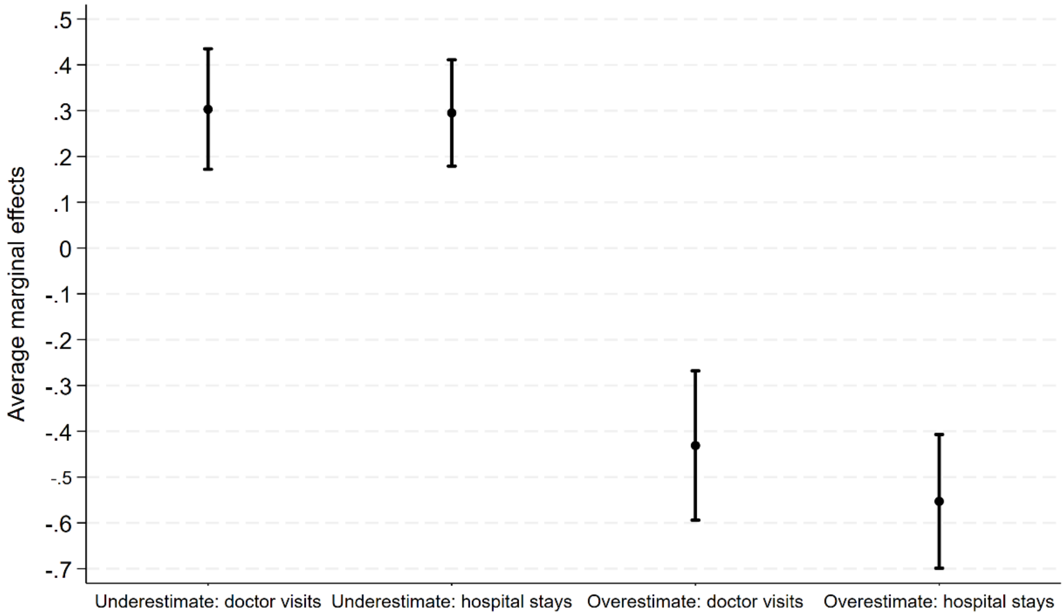

Figure. 1 Marginal effects of health perception biases on healthcare utilization (doctor visits and hospital stays)

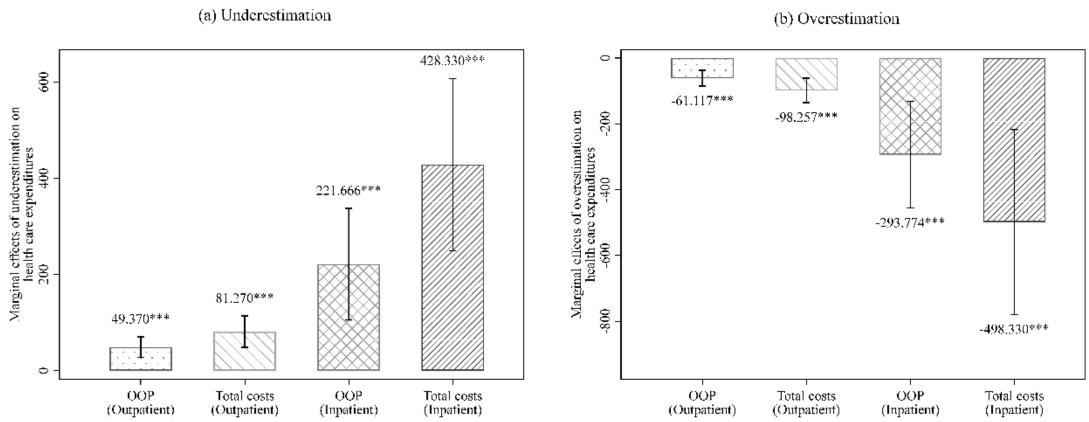

Figure. 2 Marginal effects of health perception biases on healthcare expenditures (using a gamma family and log link in the second stage of RE-TPM estimation)

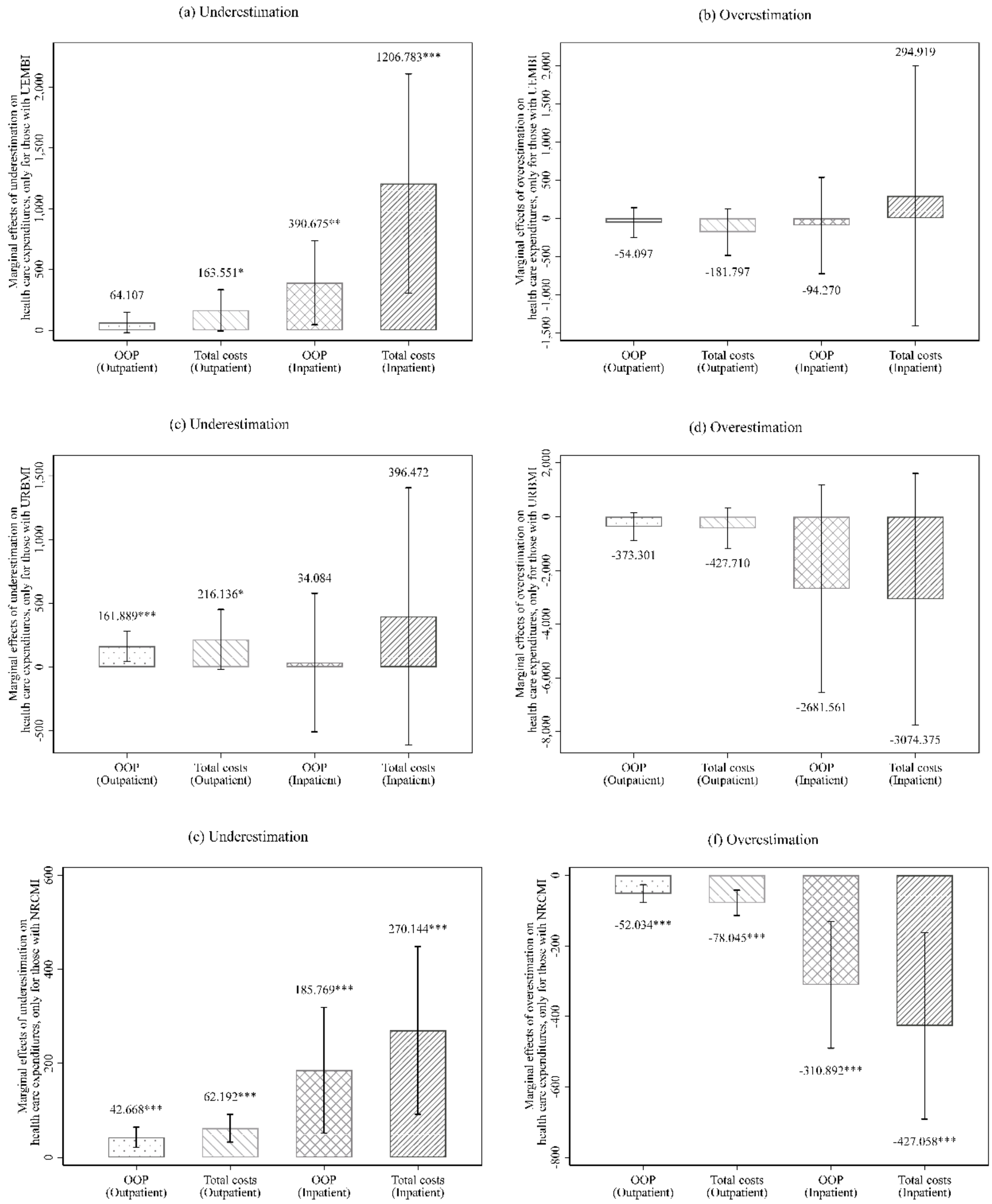

Figure. 3 Marginal effects of health misperceptions on healthcare expenditures by UEMBI, URBMI, and NRCMI (using a gamma family and log link in the second stage of RE-TPM estimation)

Note: UEBMI=the Urban Employees Basic Medical Insurance; URBMI=the Urban Residents' Basic Medical Insurance; NRCMI: the New Rural Cooperative Medical Insurance Scheme.

Appendix:

Table A1 Summary statistics for control variables: CHARLS 2011-2015

Control variables	N	Mean	S.D.	Min.	Max.	Median
Age groups						
45-54	26,431	0.304	0.460	0	1	0
55-64	26,431	0.396	0.489	0	1	0
65+	26,431	0.300	0.458	0	1	0
Male	26,431	0.479	0.500	0	1	0
Education						
Low	26,431	0.892	0.311	0	1	1
Middle	26,431	0.094	0.292	0	1	0
High	26,431	0.014	0.119	0	1	0
Married/living together	26,431	0.857	0.350	0	1	1
Retirement	26,431	0.298	0.457	0	1	0
ADLs	26,431	0.322	0.869	0	5	0
Log(household income)	26,431	8.964	2.552	0	14.929	9.674
Household size	26,431	3.374	1.707	1	15	3
Medical insurance						
No insurance	26,431	0.056	0.230	0	1	0
UEBMI	26,431	0.087	0.281	0	1	0
URBMI	26,431	0.042	0.199	0	1	0
NRCMI	26,431	0.764	0.425	0	1	1
Gongfei/PMI/others	26,431	0.052	0.222	0	1	0
Rural	26,431	0.644	0.479	0	1	1
Regions						
East	26,431	0.301	0.459	0	1	0
Center	26,431	0.296	0.457	0	1	0
West	26,431	0.327	0.469	0	1	0
Northwest	26,431	0.076	0.265	0	1	0
Survey waves						
2011	26,431	0.428	0.495	0	1	0
2013	26,431	0.310	0.462	0	1	0
2015	26,431	0.263	0.440	0	1	0

Note: ADLs=number of limitations in activities of daily living. UEBMI=Urban employee basic medical insurance, URBMI=Urban resident basic medical insurance, NRCMI=New rural cooperative medical insurance, PMI=private medical insurance. Low-level education: “at most primary school”, middle-level education: “finished middle school, high school or vocational training”, and high-level education: “two-/three-year college/associate degree, four-year college/bachelor degree, master or doctoral degree”.

Table A2 Summary statistics by health underestimation: CHARLS 2011-2015

Variables	Pos. Concordance	Mean	Underestimation	Mean	MD
Outcome variables					
Number of doctor visits	14322	0.44	3077	0.59	0.15^{***}
Number of hospital stays	14519	0.14	3126	0.24	0.09^{***}
OOP cost (outpatient)	14293	125.25	3046	220.77	95.52^{***}
Total cost (outpatient)	14355	223.27	3060	350.86	127.6
OOP cost (inpatient)	14440	598.59	3096	923.35	324.75^{***}
Total cost (inpatient)	14422	988.19	3086	1700.91	712.72^{***}
Control variables					
Age groups					
45-54	14519	0.37	3126	0.23	-0.14^{***}
55-64	14519	0.40	3126	0.43	0.03^{***}
65+	14519	0.23	3126	0.34	0.12^{***}
Male	14519	0.48	3126	0.45	-0.03^{***}
Education					
Low	14519	0.88	3126	0.90	0.02^{***}
Middle	14519	0.10	3126	0.08	-0.02^{***}
High	14519	0.01	3126	0.01	0
Married/living together	14519	0.89	3126	0.85	-0.04^{***}
Retirement	14519	0.24	3126	0.35	0.12^{***}
ADLs	14415	0.24	3120	0.42	0.18^{***}
Log(household income)	14519	9.07	3126	8.89	-0.18^{***}
Household size	14519	3.48	3126	3.19	-0.30^{***}
Medical insurance					
No insurance	14519	0.05	3126	0.05	0
UEBMI	14519	0.08	3126	0.11	0.03^{***}
URBMI	14519	0.04	3126	0.05	0.01^{**}
NRCMI	14519	0.78	3126	0.74	-0.04^{***}
Gongfei/PMI/others	14519	0.05	3126	0.06	0.01
Rural	14519	0.66	3126	0.64	-0.03^{***}
Regions					
East	14519	0.29	3126	0.32	0.02^{***}

Center	14519	0.30	3126	0.31	0.01
West	14519	0.33	3126	0.30	-0.03***
Northwest	14519	0.07	3126	0.07	0
Survey waves					
2011	14519	0.44	3126	0.36	-0.08***
2013	14519	0.30	3126	0.29	-0.01
2015	14519	0.25	3126	0.34	0.09***

Note: ADLs=number of limitations in activities of daily living. UEBMI=Urban employee basic medical insurance, URBMI=Urban resident basic medical insurance, NRCMI=New rural cooperative medical insurance, PMI=private medical insurance. MD=mean difference. Low-level education: “at most primary school”, middle-level education: “finished middle school, high school or vocational training”, and high-level education: “two-/three-year college/associate degree, four-year college/bachelor degree, master or doctoral degree”. The significance is based on independent *t*-tests. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3 Summary statistics by health overestimation: CHARLS 2011-2015

Variables	Neg. concordance	Mean	Overestimation	Mean	MD
Outcome variables					
Number of doctor visits	4824	0.55	3819	0.34	-0.21***
Number of hospital stays	4921	0.26	3865	0.13	-0.13***
OOP cost (outpatient)	4792	202.60	3805	103.44	-99.15***
Total cost (outpatient)	4816	283.63	3816	151.89	-131.74***
OOP cost (inpatient)	4866	979.89	3842	558.69	-421.21***
Total cost (inpatient)	4866	1727.82	3839	950.54	-777.28***
Control variables					
Age groups					
45-54	4921	0.20	3865	0.24	0.04***
55-64	4921	0.38	3865	0.37	0.01
65+	4921	0.42	3865	0.39	-0.03***
Male	4921	0.45	3865	0.52	0.07***
Education					
Low	4921	0.90	3865	0.91	0.01
Middle	4921	0.08	3865	0.08	0
High	4921	0.02	3865	0.01	0
Married/living together	4921	0.81	3865	0.8	-0.01
Retirement	4921	0.42	3865	0.32	-0.10***
ADLs	4907	0.51	3835	0.31	-0.20***
Log(household income)	4921	8.85	3865	8.77	-0.07
Household size	4921	3.16	3865	3.38	0.21***
Medical insurance					
No insurance	4921	0.06	3865	0.07	0.01***
UEBMI	4921	0.11	3865	0.08	-0.03***
URBMI	4921	0.05	3865	0.04	-0.01**
URCMI	4921	0.73	3865	0.76	0.02***
Gongfei/PMI/others	4921	0.05	3865	0.05	0
Rural	4921	0.6	3865	0.64	0.04***
Regions					
East	4921	0.30	3865	0.31	0

Center	4921	0.28	3865	0.28	0
West	4921	0.33	3865	0.34	0.01
Northwest	4921	0.09	3865	0.07	-0.02***
Survey waves					
2011	4921	0.39	3865	0.46	0.07***
2013	4921	0.31	3865	0.34	0.02**
2015	4921	0.29	3865	0.20	-0.09***

Note: ADLs=number of limitations in activities of daily living. UEBMI=Urban employee basic medical insurance, URBMI=Urban resident basic medical insurance, NRCMI=New rural cooperative medical insurance, PMI=private medical insurance. MD=mean difference. Low-level education= Less than lower secondary, middle-level education= Upper secondary & vocational training and high-level education= Tertiary. The significance is based on independent *t*-tests. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A4 Multinomial logistic model estimates of determinants of health underestimation and overestimation

Reference group: Concordance	Underestimation	Overestimation
Age group: 55-64	0.363***	0.270***
	(0.060)	(0.050)
Age group: 65+	0.414***	0.638***
	(0.071)	(0.057)
Male	-0.100**	0.196***
	(0.046)	(0.042)
Education: middle	-0.168*	-0.069
	(0.094)	(0.078)
Education: high	-0.158	-0.051
	(0.198)	(0.177)
Married/Partnered	0.062	-0.351***
	(0.070)	(0.060)
Retirement	0.119**	-0.031
	(0.053)	(0.050)
Number of ADLs	0.092***	-0.058**
	(0.023)	(0.023)
Log(Household income)	0.002	-0.031***
	(0.009)	(0.008)
Household size	-0.040**	0.029**
	(0.016)	(0.014)
UEBMI	0.397***	-0.296***
	(0.130)	(0.115)
URBMI	0.268*	-0.240*
	(0.149)	(0.124)
NRCMI	0.101	-0.244***
	(0.095)	(0.080)
Gongfei/PMI/Others	0.253*	-0.264**
	(0.131)	(0.109)
Rural	0.039	-0.100
	(0.072)	(0.063)

Center	-0.009	-0.105
	(0.085)	(0.075)
West	-0.173**	-0.031
	(0.081)	(0.067)
Northeast	-0.204	-0.057
	(0.124)	(0.093)
2013	0.083*	0.007
	(0.045)	(0.047)
2015	0.385***	-0.405***
	(0.046)	(0.055)
N	26277	26277
Pseudo R^2	0.019	

Note: The dependent variable is a 3-point categorical variable (1=concordance, 2=underestimation and 3=overestimation). Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Community-level clustered standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5 Random effects Poisson estimates for health perception biases on healthcare utilization (marginal effects)

	Objectively unimpaired	Objectively impaired	Objectively unimpaired	Objectively impaired
	Doctor visits		Hospital stays	
Underestimation	0.303***		0.295***	
	(0.067)		(0.059)	
95% CIs	[0.172, 0.435]		[0.179, 0.411]	
Overestimation		-0.431***		-0.553***
		(0.083)		(0.074)
95% CIs		[-0.594, -0.268]		[-0.699, -0.407]
Age group: 55-64	0.147***	0.105	0.274***	0.372***
	(0.053)	(0.095)	(0.067)	(0.132)
Age group: 65+	0.247***	0.095	0.402***	0.507***
	(0.081)	(0.098)	(0.086)	(0.141)
Gender	-0.232***	-0.241***	0.126**	0.059
	(0.062)	(0.075)	(0.062)	(0.055)
Education: middle	-0.071	-0.101	-0.166	-0.034
	(0.098)	(0.193)	(0.103)	(0.135)
Education: high	-0.013	-0.298	-0.365	-0.215
	(0.293)	(0.285)	(0.229)	(0.270)
Married/Partnered	0.029	-0.090	0.023	0.235*
	(0.087)	(0.077)	(0.076)	(0.120)
Retirement	0.331***	0.122	0.699***	0.763***
	(0.074)	(0.096)	(0.055)	(0.075)
Number of ADLs	0.164***	0.202***	0.203***	0.181***
	(0.036)	(0.033)	(0.025)	(0.036)
Household income	0.026*	-0.044***	0.002	-0.020
	(0.014)	(0.016)	(0.014)	(0.012)
Household size	0.009	0.020	-0.009	-0.013
	(0.013)	(0.023)	(0.018)	(0.026)
UEBMI	0.228	0.054	0.432**	0.622***
	(0.175)	(0.212)	(0.174)	(0.215)

URBMI	0.213	-0.397	0.647***	0.321
	(0.189)	(0.271)	(0.193)	(0.273)
NRCMI	0.181	-0.099	0.415***	0.380*
	(0.128)	(0.192)	(0.147)	(0.202)
Gongfei/PMI/Other	0.166	0.211	0.634***	0.674**
	(0.174)	(0.264)	(0.196)	(0.263)
Rural	0.211***	-0.099	0.098	-0.053
	(0.052)	(0.096)	(0.065)	(0.084)
Center	0.147***	-0.099	0.397***	0.453***
	(0.055)	(0.114)	(0.074)	(0.094)
West	0.179***	0.124	0.587***	0.630***
	(0.059)	(0.090)	(0.064)	(0.094)
Northeast	-0.632***	-0.831***	0.008	0.187
	(0.110)	(0.130)	(0.108)	(0.148)
2013	0.066	0.283***	0.372***	0.538***
	(0.052)	(0.080)	(0.070)	(0.092)
2015	-0.148**	0.052	0.395***	0.356***
	(0.069)	(0.076)	(0.060)	(0.073)
N	17289	8600	17535	8742

Note: The dependent variables are doctor visits last month and hospital stays last year. AIC= Akaike Information Criterion, BIC=Bayesian Information Criterion. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of chronic diseases, number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6 Random-effects two-part model estimates for health perception biases on healthcare expenditures (marginal effects, a gamma family and log link in the second stage)

	OOP (Outpatient) (1)	Total costs (Outpatient) (2)	OOP (Inpatient) (3)	Total costs (Inpatient) (4)
Panel A: Objectively unimpaired				
Underestimation	49.370***	81.270***	221.666***	428.330***
	(10.983)	(16.672)	(59.244)	(91.360)
N	17,231	17,307	17,426	17,398
Panel B: Objectively impaired				
Overestimation	-61.117***	-98.257***	-293.774***	-498.330***
	(12.406)	(18.953)	(82.487)	(143.435)
N	8,553	8,588	8,664	8,661

Note: The dependent variables are OOP and total costs for inpatient and outpatient, respectively. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gongfei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Robust standard errors are in parentheses. $^{*}\ p < 0.1$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table A7 Random-effects two-part model estimates for health perception biases on healthcare expenditures (marginal effects, a gaussian family and identity link in the second stage)

	OOP (Outpatient) (1)	Total costs (Outpatient) (2)	OOP (Inpatient) (3)	Total costs (Inpatient) (4)
Panel A: Objectively unimpaired				
Underestimation	67.570**	79.552	126.734*	274.568**
	(26.581)	(62.726)	(76.257)	(111.523)
N	17,231	17,307	17,426	17,398
Panel B: Objectively impaired				
Overestimation	-72.880*	-99.081**	-289.165**	-509.247***
	(37.371)	(47.912)	(125.288)	(192.006)
N	8,553	8,588	8,664	8,661

Note: Note: The dependent variables are OOP and total costs for inpatient and outpatient, respectively. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gongfei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8 Random effects Poisson estimates for health perception biases on healthcare utilization (marginal effects)

	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Panel A: Children	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UE	0.623***	0.626***			0.020	0.352**		
	(0.239)	(0.204)			(0.152)	(0.157)		
95% CIs	[0.155,1.090]	[0.226,1.027]			[-0.277,0.318]	[0.044,0.660]		
OE			-0.484**	-0.196			-0.891***	-0.701***
			(0.224)	(0.310)			(0.203)	(0.201)
95% CIs			[-0.922,-0.046]	[-0.805,0.412]			[-1.289,-0.493]	[-1.094,-0.308]
N. of child	-0.010	0.075*	-0.005	0.067	-0.013	0.031	-0.002	-0.022
	(0.036)	(0.040)	(0.029)	(0.052)	(0.032)	(0.037)	(0.040)	(0.038)
95% CIs	[-0.081,0.061]	[-0.003,0.152]	[-0.062,0.052]	[-0.036,0.170]	[-0.077,0.050]	[-0.040,0.103]	[-0.081,0.078]	[-0.097,0.052]
UE X N. of child	-0.100	-0.143**			0.059	0.042		
	(0.075)	(0.069)			(0.046)	(0.053)		
95% CIs	[-0.246,0.046]	[-0.277,-0.008]			[-0.032,0.150]	[-0.062,0.146]		
OE X N. child			-0.008	-0.040			0.056	0.129**
			(0.061)	(0.098)			(0.064)	(0.053)
95% CIs			[-0.129,0.112]	[-0.233,0.152]			[-0.070,0.181]	[0.025,0.233]
N	11425	5864	5306	3294	11551	5984	5382	3360
Panel B: Sons	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
UE	0.503***	0.585***			0.075	0.380***		
	(0.155)	(0.209)			(0.106)	(0.128)		
95% CIs	[0.199,0.806]	[0.176,0.994]			[-0.133,0.283]	[0.128,0.631]		
OE			-0.581***	-0.236			-0.969***	-0.523***
			(0.173)	(0.272)			(0.188)	(0.153)
95% CIs			[-0.920,-0.243]	[-0.769,0.297]			[-1.337,-0.601]	[-0.823,-0.222]
N. of sons	-0.014	0.071	-0.025	0.017	-0.016	-0.062	-0.114*	-0.068
	(0.051)	(0.053)	(0.042)	(0.073)	(0.037)	(0.052)	(0.060)	(0.066)
95% CIs	[-0.114,0.086]	[-0.033,0.174]	[-0.107,0.058]	[-0.127,0.161]	[-0.087,0.056]	[-0.163,0.039]	[-0.231,0.002]	[-0.198,0.062]
UE X N. of sons	-0.119	-0.257**			0.079	0.064		
	(0.089)	(0.117)			(0.065)	(0.085)		
95% CIs	[-0.294,0.056]	[-0.486,-0.027]			[-0.048,0.206]	[-0.103,0.231]		
OE X N. sons			0.045	-0.050			0.155*	0.130
			(0.080)	(0.138)			(0.080)	(0.086)
95% CIs			[-0.111,0.201]	[-0.320,0.221]			[-0.002,0.313]	[-0.038,0.298]
N	11425	5864	5306	3294	11551	5984	5382	3360
Panel C: Daut	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)

UE	0.453***	0.325**			0.159	0.440***		
	(0.159)	(0.155)			(0.126)	(0.138)		
95% CIs	[0.141,0.765]	[0.021,0.628]			[-0.088,0.405]	[0.170,0.709]		
OE			-0.440**	-0.210			-0.695***	-0.536***
			(0.175)	(0.203)			(0.148)	(0.179)
95% CIs			[-0.782,-0.097]	[-0.608,0.189]			[-0.984,-0.405]	[-0.886,-0.185]
N. of daus	0.0004	0.061	0.018	0.140*	-0.007	0.087	0.065	-0.004
	(0.039)	(0.062)	(0.037)	(0.072)	(0.037)	(0.056)	(0.053)	(0.057)
95% CIs	[-0.077,0.077]	[-0.061,0.182]	[-0.055,0.091]	[-0.002,0.281]	[-0.080,0.065]	[-0.022,0.196]	[-0.039,0.170]	[-0.116,0.107]
UE X N. of daut	-0.081	-0.056			0.028	0.017		
	(0.081)	(0.114)			(0.065)	(0.084)		
95% CIs	[-0.239,0.078]	[-0.279,0.167]			[-0.100,0.157]	[-0.147,0.181]		
OE X N. daut			-0.047	-0.073			-0.006	0.139*
			(0.091)	(0.124)			(0.106)	(0.081)
95% CIs			[-0.226,0.132]	[-0.316,0.170]			[-0.214,0.201]	[-0.020,0.297]
N	11425	5864	5306	3294	11551	5984	5382	3360

Note: UE=underestimation; OE=overestimation. The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender, education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9 Random-effects two-part model estimates for health perception biases on healthcare expenditures (marginal effects)

	OOP (Outpatient) (1)	Total costs (Outpatient) (2)	OOP (Inpatient) (3)	Total costs (Inpatient) (4)
Panel A: UEMBI				
Objectively unimpaired				
Underestimation	64.107	163.551*	390.675**	1,206.783***
	(43.293)	(86.533)	(176.379)	(459.755)
N	1,387	1,421	1,434	1,429
Objectively impaired				
Overestimation	-54.097	-181.797	-94.270	294.919
	(99.334)	(155.776)	(322.444)	(869.301)
N	807	820	829	828
Panel B: UEMBI				
Objectively unimpaired				
Underestimation	161.889***	216.136*	34.084	396.472
	(60.490)	(119.880)	(277.081)	(514.787)
N	663	667	672	671
Objectively impaired				
Overestimation	-373.301	-427.710	-2,681.561	-3,074.375
	(262.152)	(386.404)	(1,970.150)	(2,390.238)
N	409	410	418	417
Panel C: NRCMI				
Objectively unimpaired				
Underestimation	42.668***	62.192***	185.769***	270.144***
	(11.022)	(15.069)	(68.131)	(91.000)
N	13,386	13,407	13,497	13,471
Objectively impaired				
Overestimation	-52.034***	-78.045***	-310.892***	-427.058***
	(12.884)	(18.368)	(91.528)	(134.848)
N	6,384	6,392	6,439	6,437

Note: The dependent variables are OOP and total costs for inpatient and outpatient, respectively. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gongfei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Robust standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A10 Random effects Poisson estimates for health perception biases on healthcare utilization by gender (marginal effects)

Panel A: Males	Doctor visits		Hospital stays	
Underestimation	0.393***		0.382***	
	(0.109)		(0.094)	
95% CIs	[0.179, 0.607]		[0.197, 0.567]	
Overestimation		-0.483***		-0.658***
		(0.118)		(0.117)
95% CIs		[-0.714, -0.252]		[-0.887, -0.428]
N	8282	4152	8393	4209
Panel B: Females				
Underestimation	0.232***		0.218***	
	(0.086)		(0.083)	
95% CIs	[0.063, 0.400]		[0.056, 0.380]	
Overestimation		-0.408***		-0.473***
		(0.107)		(0.123)
95% CIs		[-0.618, -0.199]		[-0.714, -0.233]
P-value (b_1=b_2)	0.125	0.553	0.124	0.000
N	9007	4448	9142	4533

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11 Random effects Poisson estimates for health perception biases on healthcare utilization by age groups (marginal effects)

Panel A: Aged 45-54	Doctor visits		Hospital stays	
Underestimation	0.343^{***}		0.380^{**}	
	(0.118)		(0.157)	
95% CIs	[0.112, 0.574]		[0.071, 0.688]	
Overestimation		-0.490^{***}		-0.753^{***}
		(0.187)		(0.203)
95% CIs		[-0.856, -0.124]		[-1.151, -0.354]
N	6001	1845	6074	1866
Panel B: Aged 55-64				
Underestimation	0.423^{***}		0.221^{**}	
	(0.095)		(0.098)	
95% CIs	[0.238, 0.609]		[0.029, 0.413]	
Overestimation		-0.459^{***}		-0.619^{***}
		(0.114)		(0.131)
95% CIs		[-0.682, -0.237]		[-0.875, -0.362]
N	7025	3238	7122	3300
Panel C: Aged 65+				
Underestimation	0.162		0.356^{***}	
	(0.113)		(0.088)	
95% CIs	[-0.061, -0.384]		[0.183, 0.529]	
Overestimation		-0.405^{***}		-0.460^{***}
		(0.121)		(0.138)
95% CIs		[-0.642, -0.167]		[-0.730, -0.189]
P-value (b_1=b_2)	0.500	0.923	0.231	0.144
P-value (b_2=b_3)	0.048	0.735	0.312	0.220
N	4263	3517	4339	3576

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12 Random effects Poisson estimates for health perception biases on healthcare utilization (marginal effects)#

Panel A: Nonpoor	Doctor visits		Hospital stays	
Underestimation	0.228***		0.285***	
	(0.083)		(0.089)	
95% CIs	[0.065, 0.391]		[0.110, 0.460]	
Overestimation		-0.333***		-0.445***
		(0.127)		(0.118)
95% CIs		[-0.582, -0.084]		[-0.676, -0.215]
N	8871	4027	9013	4099
Panel B: Poor				
Underestimation	0.376***		0.264***	
	(0.097)		(0.090)	
95% CIs	[0.186, 0.567]		[0.087, 0.440]	
Overestimation		-0.530***		-0.639***
		(0.094)		(0.106)
95% CIs		[-0.714, -0.347]		[-0.845, -0.432]
P-value (b_1=b_2)	0.000	0.225	0.765	0.104
N	8418	4573	8522	4643

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.
The poor dummy equal to 1 if per capita household income in our analytic sample is equal or below its median level, 0 otherwise.

Table A13 Random effects Poisson estimates for health perception biases on healthcare utilization (marginal effects)

Panel A: Rural	Doctor visits		Hospital stays	
Underestimation	0.339***		0.198***	
	(0.087)		(0.070)	
95% CIs	[0.168,0.509]		[0.061,0.336]	
Overestimation		-0.507***		-0.702***
		(0.110)		(0.104)
95% CIs		[-0.723, -0.292]		[-0.905, -0.498]
N	11425	5306	11551	5382
Panel B: Urban				
Underestimation	0.256**		0.463***	
	(0.126)		(0.085)	
95% CIs	[0.008, 0.503]		[0.297, 0.629]	
Overestimation		-0.300*		-0.343***
		(0.158)		(0.112)
95% CIs		[-0.608, 0.009]		[-0.562, -0.124]
P-value (b_1=b_2)	0.221	0.463	0.000	0.001
N	5864	3294	5984	3360

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender, education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14 Random effects Poisson estimates for health perception biases in 2011 on healthcare utilization in 2015 (marginal effects)

	Doctor visits		Hospital stays	
Underestimation	0.073		0.382*	
	(0.155)		(0.207)	
95% CIs	[-0.230, 0.376]		[-0.024, -0.787]	
Overestimation		-0.461*		-0.577***
		(0.245)		(0.186)
95% CIs		[-0.942, -0.020]		[-0.941, -0.214]
N	6251	2957	6337	3006

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A15 Fixed effects Poisson/fixed effects negative binomial/random effects negative binomial estimates for health perception biases on healthcare use (marginal effects)

	Doctor visits		Hospital stays	
Panel A: Fixed effects poission				
Underestimation	0.624***		0.016	
	(0.113)		(0.226)	
95% CIs	[0.402, 0.846]		[-0.427, 0.459]	
Overestimation		-0.410***		-0.942***
		(0.152)		(0.274)
95% CIs		[-0.708, -0.112]		[-1.479, -0.405]
N	4405	1533	2590	1144
Panel B: Fixed effects negative binomial				
Underestimation	0.305*		-0.071	
	(0.164)		(0.206)	
95% CIs	[-0.015, 0.626]		[-0.475, 0.333]	
Overestimation		-0.421**		-0.718**
		(0.181)		(0.289)
95% CIs		[-0.776, -0.065]		[-1.285, -0.151]
N	4405	1533	2590	1144
Panel C: Random effects negative binomial				
Underestimation	0.300***		0.341***	
	(0.043)		(0.059)	
95% CIs	[0.215, 0.385]		[0.225, 0.456]	
Overestimation		-0.491***		-0.532***
		(0.061)		(0.074)
95% CIs		[-0.609, -0.372]		[-0.677, -0.388]
N	17289	8600	17535	8742

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A16 Random effects probit estimates for different outpatient/inpatient behaviors on health perception biases (marginal effects)

	Underestimation	Overestimation
Panel A: A follow-up vs the first visit		
For last outpatient: a follow-up visit	0.002	-0.097^{***}
	(0.004)	(0.035)
95% CIs	[-0.006, 0.010]	[-2.139, -0.133]
N	3363	1643
Panel B: Purpose for last outpatient visit		
Medical check-ups	0.044	-0.142
	(0.043)	(0.345)
95% CIs	[-0.040, 0.129]	[-0.818, 0.534]
N	3658	1809

Note: The dependent variables are health underestimation or overestimation. Controls include a dummy for the first visit for last outpatient/purposes for last outpatient visit (1=others, 2=immunization/consultation, 3=medical check-ups, and 4= treatment for illness, 1=others as the reference group), age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A17 Random effects Poisson/generalized linear/latent class Poisson estimates for health perception biases on healthcare utilization

	Doctor visits		Hospital visits	
Panel A: Latent class Poisson model				
Underestimation	0.215***		0.281***	
	(0.027)		(0.044)	
95% CIs	[0.162, 0.269]		[0.195, 0.367]	
Overestimation		-0.418***		-0.521***
		(0.035)		(0.054)
95% CIs		[-0.486, -0.350]		[-0.626, -0.416]
N	17289	8600	17535	8742
Panel B: Multilevel mixed-effects Poisson regression (marginal effects)				
Underestimation	0.433***		0.597***	
	(0.081)		(0.095)	
95% CIs	[0.273, 0.592]		[0.410, 0.783]	
Overestimation		-0.881***		-0.740***
		(0.118)		(0.136)
95% CIs		[-1.113, -0.650]		[-1.007, -0.472]
N	17289	8600	17535	8742

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A18 Random effects Poisson/generalized linear/latent class Poisson estimates for health perception biases on healthcare utilization

	Doctor visits		Hospital visits	
Panel A: Random-effects Poisson model with cluster-robust standard errors (marginal effects)				
Underestimation	0.301***		0.280***	
	(0.069)		(0.059)	
95% CIs	[0.165, 0.437]		[0.164, 0.395]	
Overestimation		-0.422***		-0.572***
		(0.082)		(0.080)
95% CIs		[-0.583, -0.261]		[-0.730, -0.415]
N	17289	8600	17535	8742
Panel B: Generalized linear model (GLM) with heteroskedasticity-consistent variance				
Underestimation	0.215***		0.281***	
	(0.058)		(0.058)	
95% CIs	[0.101, 0.329]		[0.168, 0.394]	
Overestimation		-0.418***		-0.521***
		(0.078)		(0.082)
95% CIs		[-0.571, -0.265]		[-0.681, -0.360]
N	17289	8600	17535	8742

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A19 Random effects Poisson estimates for health perception biases on healthcare utilization (using dyslipidemia and diabetes, marginal effects)

	Doctor visits		Hospital stays	
Panel A: Health perception based on dyslipidemia				
Underestimation	0.376***		0.478***	
	(0.087)		(0.094)	
95% CIs	[0.206, 0.547]		[0.294, 0.663]	
Overestimation		-0.379***		-0.620***
		(0.118)		(0.093)
95% CIs		[-0.611, -0.148]		[-0.802, -0.437]
N	9290	5553	9403	5668
Panel B: Health perception based on diabetes				
Underestimation	0.554***		0.602***	
	(0.133)		(0.114)	
95% CIs	[0.293, 0.814]		[0.379, 0.825]	
Overestimation		-0.587***		-0.539***
		(0.121)		(0.173)
95% CIs		[-0.830, -0.350]		[-0.879, -0.199]
N	11595	1580	11763	1609

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A20 Random effects Poisson estimates for health perception biases on healthcare utilization (control for SAH, marginal effects)

	Doctor visits		Hospital stays	
Underestimation	0.223***		0.227***	
	(0.065)		(0.060)	
95% CIs	[0.095, 0.351]		[0.109, 0.344]	
Overestimation		-0.331***		-0.442***
		(0.076)		(0.096)
95% CIs		[-0.481, -0.181]		[-0.630, -0.254]
SAH: Fair	1.494***	2.090***	1.714***	2.261***
	(0.152)	(0.275)	(0.169)	(0.249)
SAH: Poor	2.989***	4.093***	4.093***	4.946***
	(0.349)	(0.537)	(0.444)	(0.520)
N	17384	8634	17630	8777

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), self-assessed health (SAH, 1=excellent/good, 2=fair and 3=poor, with 1=excellent/good as the reference group), translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A21 Random effects Poisson estimates for health perception biases on healthcare utilization (marginal effects)

	Doctor visits		Hospital visits	
Panel A: Variable addition test				
Underestimation	0.268***		0.330***	
	(0.071)		(0.076)	
95% CIs	[0.129, 0.406]		[0.180, 0.479]	
Overestimation		-0.406***		-0.483***
		(0.095)		(0.101)
95% CIs		[-0.592, -0.220]		[-0.682, -0.285]
Number of	0.058	-0.061	-0.104	-0.119
	(0.055)	(0.072)	(0.072)	(0.092)
N	17289	8600	17535	8742
Panel B: Balanced panel				
Underestimation	0.333***		0.203**	
	(0.094)		(0.098)	
95% CIs	[0.149, 0.516]		[0.011, 0.396]	
Overestimation		-0.315**		-0.542***
		(0.152)		(0.119)
95% CIs		[-0.613, -0.017]		[-0.775, -0.309]
N	6251	2957	6337	3006

Note: The dependent variables are doctor visits last month and hospital stays last year. Controls include age groups (45-54, 55-64 and 64+, 45-54 as the reference group), gender (1=male, 0=female), education level (1=less than lower secondary, 2=upper secondary and vocational training and 3=tertiary, with 1=less than lower secondary as the reference group), marital status (1=married/partnered, 0=otherwise), retirement (1=yes, 0=no), number of ADLs, translog household income, household size, medical insurance (1=no insurance, 2=UEBMI, 3=URBMI, 4=NRCMI and 5=Gong-fei/PMI/others, with 1=no insurance as the reference group), rural, wave and regional dummies (1=east, 2=center, 3=west and 4=northwest, with 1=east as the reference region). Bootstrap standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

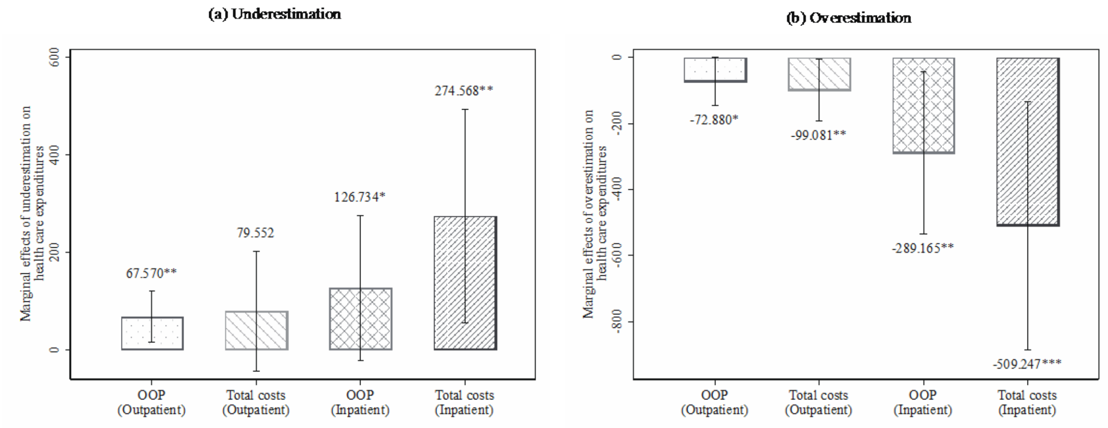

Figure. A1 Marginal effects of health perception biases on healthcare expenditures (using a gaussian a gamma family and log link in the second stage of RE-TPM estimation)