Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Quantified self and healthy food consumption: evidence from a household survey in China

**Yaling Zhu**
Shaanxi Normal University
**Jinyu Chen**
Shaanxi Normal University
**Xiangguo Zhang**
xiangguozhang@mail.tsinghua.edu.cn

Tsinghua University

**Article**

**Keywords:** Food consumption, Healthy food, Quantified self, Self-tracking motivation, Household food expenditure, China

**Posted Date:** November 17th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11117318/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Quantified self and healthy food consumption: evidence from a household survey in China

**Abstract** – As data-driven technologies become increasingly embedded in everyday life, quantified-self practices are transforming how consumers track and manage their health. This study investigates whether engagement in the quantified self is associated with healthy food consumption among urban households in China. Using household survey data collected from twelve cities across six provinces, ordinary least squares models are estimated with a rich set of controls and occupation-city fixed effects, supplemented by instrumental-variable estimation to address potential endogeneity. The results show that greater engagement in the quantified self is positively associated with a higher healthy-food expenditure share. These findings remain robust in the instrumental-variable analysis. Further analysis indicates that self-optimisation and self-supervision motivations strengthen the relationship between quantified-self engagement and the healthy-food expenditure share, whereas self-enjoyment motivation does not. Heterogeneity tests across twenty-two potential moderators suggest that the association is broadly stable rather than concentrated in particular subgroups, and only the groups defined by smoking status exhibit statistically significant between-group differences. Overall, the findings highlight the role of self-tracking practices in shaping healthier food-related preferences and spending behaviour in urban China.

**Keywords:** Food consumption; Healthy food; Quantified self; Self-tracking motivation; Household food expenditure; China

## Introduction

Rising living standards and household incomes, together with the global rise in obesity and other noncommunicable diseases, have increased consumers' attention to diet. Expectations about food consumption have shifted from meeting basic calorie needs to emphasising health, nutrition, and safety, consistent with growing demand for healthy foods. A market report by Research and Markets [1] indicates that the global health food market exhibited a steady expansion between 2015 and 2024. In the Chinese context, Daxue Consulting highlights that

[1] Research and Markets :Health and Wellness Market Size, Share, and Growth Analysis, By Dietary Supplements (Vitamins & Minerals, Herbal Supplements), By Weight Loss, By Functional Foods, By Region -Industry Forecast 2026-2033, published through GII.

https://www.gii.tw/report/sky1902887-health-wellness-market-size-share-growth-analysis.html

the country's health and wellness sector ranks as the second largest worldwide, with an estimated market value of USD 950 billion. Furthermore, the protein market in China is projected to grow from USD 2.76 billion in 2025 to USD 2.96 billion in 2026[2]. Consistent with these trends, both health awareness and healthy food consumption among Chinese residents have increased.

Greater attention to dietary health has coincided with improvements in overall nutrition. Consistent with this pattern, stunting and malnutrition among Chinese children declined from 17.8% and 10.0% in 2000 to 4.8% and below 2.5% in 2022, respectively (Lin and Hsu 2015). Despite these gains, dietary intake remains structurally imbalanced. Current patterns include high consumption of meat and fat, low intake of cereals, and widespread vitamin deficiencies (Lin and Hsu 2015). A central factor is limited understanding of nutritional content and inadequate assessment of dietary appropriateness. Better knowledge of nutritional components, together with food choices aligned with individual nutritional needs, could help improve diet quality and mitigate both undernutrition and overnutrition in China(Cheng et al. 2025).

China's dietary transition is increasingly digital. Internet penetration has been shown to reshape rural diets and improve dietary quality (Gao et al. 2026), digital technology more broadly reshapes household production and consumption decisions (Hou et al. 2019), and online information channels now mediate how consumers learn about nutrition and food safety (Yang et al. 2020). Within this broader digitalisation of food choice, quantified-self practices—the self-directed tracking of personal health and dietary data through apps, wearables, and mini-programs—represent the consumer-level micro-foundation of digital nutrition.

Quantified self refers to monitoring and measuring individuals' activities to develop self perception and regulate daily routines (Fan et al. 2022; Robertsson 2014; Rooksby et al. 2014). By providing structured feedback on behaviour, quantified self can help consumers better understand their health status and adjust decisions in domains such as work and consumption (Guo 2016; Shen and Li 2017). Because quantified self is intended to improve health outcomes and healthy food consumption is a key indicator of nutritional status, quantified self may affect the composition of food consumption (Lei et al. 2021).

Although a growing literature studies quantified self in the context of healthy food consumption, the evidence is mixed. From an information perception perspective, quantification translates

[2] https://daxueconsulting.com/protein-market-china/

technical information, such as food safety assessments, into data that consumers can interpret (Ruckenstein and Pantzar 2017). In this view, quantified self behaviour may lead consumers to choose products based on information about ingredients and safety, but these responses can be short-lived and do not necessarily improve consumers' nutritional understanding or longer-run health behaviour (Gupta et al. 2020; Yang et al. 2020; Zhang et al. 2020); a recent meta-analysis likewise finds heterogeneous effects of self-quantification on consumer well-being (Jain et al. 2025). In contrast, risk perception theory emphasises that, when individuals perceive health risks, they collect, manage, and analyse quantified information about their own health and needs, such as blood pressure and blood glucose, which can support more targeted behavioural adjustments(Almalki et al. 2016). This mechanism relies on individual health status to guide food choice and implies a positive role for quantified self in shaping nutrition concepts and healthy behaviours(Almalki et al. 2016; Fan et al. 2022).

The divergence in prior findings motivates further study of the role of quantified self in healthy food consumption. It remains unclear whether quantified self improves health-related decision-making or plays only a limited role in shaping nutrition concepts; differences may reflect heterogeneity in motives and tracking intensity, and incomplete treatment of endogeneity. We therefore analyse this association using survey data on 1,454 individuals from 12 cities across 6 provinces of China, testing whether quantified self is associated with a greater desire for healthy foods and a larger healthy-food expenditure share after addressing endogeneity, and whether it varies by motives and consumer characteristics. The paper makes three contributions. First, it provides household-level evidence linking quantified-self engagement to the nutritional composition of food demand in China through a money-metric outcome—the healthy-food expenditure share—connecting the quantified-self literature to household food-demand economics, where recent work on digitalisation and diets has focused on internet access rather than consumer self-tracking (Gao et al. 2026). Second, it shows that the association varies systematically with quantified-self motivations—stronger for self-optimisation and self-supervision than for self-enjoyment—linking the motivational typology of self-tracking (Li and Zhang 2018; Ruckenstein and Pantzar 2017; Zhang et al. 2019) to food-consumption outcomes. Third, it documents heterogeneity along consumer, family, and technology dimensions, identifying the groups for whom digital self-monitoring is most complementary to information-based food policy.

## Theoretical analysis and hypothesis formulation

### Quantified self and residents' healthy food consumption

Quantified self involves tracking consumers' physiological and behavioural information by recording aspects of daily life and generating statistical indicators of habits and behaviours. Greater attention to quantified information about food may increase consumers' sensitivity to potentially harmful components, improving their ability to assess food safety risks and perceived efficacy(Firoozzare et al. 2024; Yang et al. 2020). As residents move from an intention to engage in quantified self to actual tracking behaviour, their degree of engagement may increase over time, strengthening the desire to consume healthy food and translating into specific behaviours. Consumers may adjust decisions using information derived from tracking and interpreting self-related data, such as the maximum tolerable intake of specific additives or their own daily consumption, together with product information such as additive content (Zhang et al. 2020). In response to perceived health risks, residents may also invest more effort in collecting, managing, and analysing quantified information about their own health, such as blood pressure and blood glucose, supporting more targeted behavioural adjustments (Schroeder et al. 2007). With clearer feedback, residents may rely more on quantified information about their own needs rather than on subjective judgment (Almalki et al. 2016). Thus, we propose the following hypothesis:

Hypothesis 1: Engagement in quantified self is positively associated with the healthy-food expenditure share.

### Motivations for quantified self on residents' healthy food consumption

Residents' participation in quantified self is primarily driven by three motivations: self-optimisation motivation, self-supervision motivation, and self-enjoyment motivation (Kampen 2020; Li and Zhang 2018). Self-optimisation motivation leads residents to engage in quantified self to improve outcomes and adjust behaviour. Relative to participation without a specific goal, residents with stronger self-optimisation motivation may be more likely to engage in quantified self and to purchase healthy foods based on the resulting information in order to improve health or reduce illness risk (Fan et al. 2022). Residents may rely more on quantified information in food consumption, including past records of physiological and behavioural measures and information about product attributes, making consumption decisions more targeted.

Quantified self under self-supervision motivation is intended to strengthen self-control (Fan et al. 2022). Prior work shows that self-tracking can support monitoring and self-regulation by

giving users structured, data-driven feedback on their own behaviour (Ruckenstein and Pantzar 2017; Shin and Biocca 2017). With stronger self-supervision motivation, residents may place greater trust in quantified self data to discipline behaviour, such as reducing tobacco and alcohol intake, limiting sugary foods to manage blood glucose, and seeking information that supports consistent self monitoring.

Self-enjoyment motivation reflects consumers' interest in related activities and devices, which can increase participation in quantified self and, in turn, influence consumption decisions. By recording data accurately and continuously, quantified self can increase residents' understanding of dietary intake and health status, which may increase willingness to adjust food choices to improve diet quality(Zhang et al. 2020). Under self-enjoyment motivation, residents may also pursue titles (e.g., environmental guardian or vegetarian) or rewards (e.g., a series of products such as reusable bags or tableware). These features may encourage continued self tracking and increase reliance on self quantified data when making food consumption decisions. Based on this discussion, we propose the following hypothesis:

Hypothesis 2: Motivations moderate the association between quantified-self engagement and healthy food consumption: self-optimisation (3a) and self-supervision (3b) strengthen the positive association.

Motivations may differ in the extent to which they sustain continued participation and the search for ongoing feedback. Self-optimisation and self-supervision motivations emphasise improving behavioural processes and personal capabilities (Swan 2009), whereas self-enjoyment motivation emphasises outcomes and achievements. Relative to the former, self-enjoyment may involve lower cognitive burden and higher enjoyment of participation, which can strengthen motivation in the short run but may be more susceptible to interruption if rewards become costly or interest declines. In contrast, self-optimisation and self-supervision motivations reflect more persistent needs: by using tracking data, residents receive repeated feedback and may treat the results as diagnostic information about how to improve behaviour (Swan 2009). Residents motivated primarily by self-enjoyment, by contrast, may be more likely to stop seeking feedback and respond by shifting goals or disengaging (Liu et al. 2013). Based on this reasoning, we propose the following hypothesis:

Hypothesis 3: The moderating effect of self-enjoyment motivation is weaker than that of self-optimisation and self-supervision motivations.

# Methods

## Survey and sample

Based on existing scales and interviews with experts, we developed measurement scales for the key variables, and finalised the questionnaire after a pretest with five scholars and 30 consumers. To improve coverage, we administered it both online and in person to urban and rural residents and to college students. During the field period, we distributed 1,600 questionnaires and received 1,501 responses, covering 12 cities in Zhejiang, Jiangsu, Henan, Shanxi, Shaanxi, and Gansu provinces. The sample covers China's eastern, central and western regions, captures urban-rural heterogeneity, spans a multi-tiered economic gradient, and represents diverse regional dietary cultures, which controls for potential confounding effects of cultural background. We excluded 47 responses as invalid because of missing items or implausible response patterns. After excluding key missing values and outliers, the final sample includes 1,454 valid questionnaires, for an effective response rate of 90.9%. Table 1 reports the demographic characteristics of the respondents.

**Table 1.** Demographic characteristics of respondents

| Category | | Number | Percentage | Category | | Number | Percentage |
|---|---|---|---|---|---|---|---|
| Gender | Male | 664 | 45.67 | Residence | urban | 575 | 39.55 |
| | Female | 790 | 54.33 | | rural | 879 | 60.45 |
| Age | 25 and below | 79 | 5.43 | Province | Zhejiang | 308 | 21.18 |
| | 26-35 | 553 | 38.03 | | Jiangsu | 169 | 11.62 |
| | 36-45 | 460 | 31.64 | | Henan | 324 | 22.28 |
| | 46 and above | 362 | 24.90 | | Shanxi | 280 | 19.26 |
| Monthly income per capita | 2,000 yuan and less | 302 | 20.77 | | Shaanxi | 232 | 15.96 |
| | 2,001-5,000 yuan | 958 | 65.89 | | Gansu | 141 | 9.70 |
| | 5001-10000 yuan | 172 | 11.83 | Source | Online | 614 | 42.23 |
| | More than 10,000 yuan | 22 | 1.51 | | Community | 462 | 31.77 |
| Education | Below bachelor degree | 1198 | 82.39 | | Park | 225 | 15.47 |
| | Bachelor degree and above | 256 | 17.61 | | School | 153 | 10.52 |

## Variables and measurement

### *Dependent variables*

### Healthy food consumption desire

Drawing on the literature on Chinese consumers' food-safety attitudes (Liu et al. 2013; Schroeder et al. 2007) and on awareness and trust in healthy-food purchases (Firoozzare et al. 2024), we measure residents' desire to consume healthy food with five items covering attention to food-nutrition and food-safety information, scrutiny of quality-inspection labels and

packaging content, concern about the health effects of unsafe food, and willingness to purchase healthy, organic and safe food; the items are reproduced in Supplementary Note 1 (Block 1). All items were rated on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). They tap distinct facets of food-safety and nutrition concern and are only weakly intercorrelated (mean inter-item correlation 0.13, Cronbach's alpha 0.43), so we treat them as a composite index of stated desire rather than a unidimensional reflective scale. The index is used only in the robustness analysis (Table 3, column (7)); all main results rest on the money-metric expenditure share.

**Healthy food expenditure share**

We measure healthy food consumption as the share of the household's monthly expenditure on health-oriented food products in total monthly food expenditure—a money-metric measure of dietary composition that is standard in household food-demand research (Castro Campos and Qi 2024). The definition of health-oriented food products shown to respondents follows the healthy-food categories used in recent Chinese consumer studies (Firoozzare et al. 2024); the full wording appears in Supplementary Note 1.

***Independent variable***

**(1) Degree of quantified self**

Drawing on the personal-informatics and self-tracking literature (Almalki et al. 2016; Ruckenstein and Pantzar 2017; Shin and Biocca 2017) and the Chinese quantified-self literature (Li and Zhang 2018), we measure engagement in quantified self with three items: "I am concerned about changes in my health indicators," "I am willing to make rational decisions about food consumption by evaluating my health indicators," and "I am willing to select alternative safe food products with equivalent nutritional efficacy based on my self-quantified dietary needs." The three items load on a single factor in a confirmatory factor analysis; because a one-factor model with three indicators is just identified, no global fit indices are available. Internal consistency is acceptable (Cronbach's alpha = 0.652). Etkin (2016) and Zhang et al. (2019) provide the theoretical motivation for the construct but are not the source of the items.

***Moderators***

**(1) Quantified self motivations**

To examine whether motivations moderate the relationship between quantified self and healthy food consumption, we measure three motivations—self-optimisation, self-supervision, and self-enjoyment—following the typology in the Chinese quantified-self literature (Li and Zhang

2018), the goal-pursuit account of quantification's motivational effects (Zhang et al. 2019), and the dataistic-paradigm account of sustained tracking (Ruckenstein and Pantzar 2017). The nine-item instrument is reproduced in Supplementary Note 1 (Block 4). The analysis dataset provides one five-point indicator per motivation, and we use these three indicators directly as moderators; they are therefore summary indicators of motivational orientation rather than multi-item scale scores.

***Control variables***

Referring to Kersten-van Dijk et al. (2017), Li and Zhang (2018) and Biroli et al. (2022), we include hukou registration status, age, gender, education, marital status, self-reported health status, the logarithm of monthly household income, and health-insurance coverage as the main control variables, following the consumption literature that treats income as a primary determinant of household spending (Ye et al. 2011), together with occupation and city fixed effects.

***Heterogeneity dimensions***

Following Zhang et al. (2019) and Biroli et al. (2022), we conduct heterogeneity analyses along four dimensions: individual characteristics (age, gender, education, self-rated health, income, and health-insurance coverage), household characteristics (number of elderly members, number of young children, number of household members in poor health, health-management needs, and the family social network), quantified-self-related characteristics (attitudes toward health technology, experience with quantitative techniques, digital ability, trust and privacy concerns, and the duration and frequency of quantified self), and food-and-health-related characteristics (health awareness, food-safety awareness, food-nutrition awareness, drinking, and smoking). For each moderator we report the marginal effect of quantified-self engagement at a low and a high value, together with a test of whether the two marginal effects are equal. Descriptive statistics for all variables are reported in Supplementary Table S5.

**Empirical methods and models**

Based on the data and the relationships among the variables, we estimate the association between quantified self and healthy food consumption using the following econometric model:

$$y_i = \alpha_{1i} quantified_i + \alpha_{2i} X_i + \varepsilon_i \quad (1)$$

In Equation (1), $y_i$ denotes the dependent variable (healthy food consumption desire or the healthy-food expenditure share), $quantified_i$ is the degree of quantified self, and $X_i$ is a vector of control variables. Our coefficient of interest is $\alpha_{1i}$; a positive estimate indicates that a higher

degree of quantified self is associated with greater healthy food consumption. Occupation fixed effects refer to the respondent's reported occupation or industry sector; because quantified-self use and food expenditure patterns vary systematically across occupations, all specifications include occupation fixed effects, which in our individual-level cross-section absorb occupational differences in income, digital exposure, and dietary patterns.

To examine whether quantified self motivations moderate the relationship between quantified self and healthy food consumption, we estimate the following specification:

$$y_i = \beta_{1i} Quantified_i + \beta_{2i} Motive + \beta_{3i} Motive \times Quantified_i + \beta_{4i} X_i + \varepsilon_i \quad (2)$$

We also assess heterogeneity by estimating:

$$y_i = \beta_{1i} Quantified_i + \beta_{2i} Inter + \beta_{3i} Inter \times Quantified_i + \beta_{4i} X_i + \varepsilon_i \quad (3)$$

In Equation (2) and (3), *Motive* denotes the quantified self-motivation measures (self-optimisation, self-supervision, and self-enjoyment) and *Inter* denotes the heterogeneity variables used in the subgroup and interaction analyses.

Reverse causality is a first-order concern, because consumers who already prefer healthy food may be more likely to monitor their diet in quantitative terms, and the survey is clustered by city and by age cohort. We therefore complement the ordinary-least-squares estimates with an instrumental-variables specification (Kuhlgatz et al. 2023) and report inference that is robust to the survey design; the identification argument, the test statistics and the full heterogeneity results are given in the endogeneity and robustness analysis and the heterogeneity analysis subsections.

## Results

### Benchmark estimates

We first estimate the relationship between quantified self and healthy food consumption; the results are reported in Table 2. Columns (1) and (2) report specifications with fixed effects only and with the full set of controls, both using the share of health-oriented food in total food expenditure as the dependent variable. The coefficient on quantified self is 0.129 (standard error 0.004) without controls and 0.128 (0.004) with controls, both significant at the 1% level. Column (7) of Table 3 reports the corresponding estimate for stated desire (1.484, 0.048). A one standard deviation increase in quantified-self engagement is associated with a 0.64 standard deviation increase in the healthy-food expenditure share and a 0.63 standard deviation increase in stated desire.

**Table 2.** Baseline association between quantified self and healthy food consumption

| | (1)<br>share | (2)<br>share |
|---|---|---|
| selfqua | 0.129*** | 0.128*** |
| | (0.004) | (0.004) |
| register | | 0.002 |
| | | (0.007) |
| age | | -0.000 |
| | | (0.000) |
| gender | | -0.003 |
| | | (0.007) |
| edu | | -0.007*** |
| | | (0.002) |
| marry | | -0.005 |
| | | (0.006) |
| health | | -0.001 |
| | | (0.002) |
| income | | 0.003 |
| | | (0.002) |
| insur | | 0.014* |
| | | (0.008) |
| constant | 0.245*** | 0.254*** |
| | (0.010) | (0.027) |
| Occupation fixed effects | Yes | Yes |
| City fixed effects | Yes | Yes |
| N | 1454 | 1454 |
| adj. $R^2$ | 0.414 | 0.416 |
| F statistic | 40.45 | 31.49 |

Notes: Robust standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. This convention applies to all subsequent tables.

Education is the only control statistically significantly associated with the healthy-food expenditure share (−0.007, standard error 0.002), so that more educated respondents allocate a smaller share of food expenditure to health-oriented products. Health-insurance coverage enters positively at the 10% level (0.014, 0.007); hukou registration, age, gender, marital status, health status, and log income are not individually significant. In the desire equation (Table 3, Column (7)), women report significantly higher desire than men (0.164, 0.076) and respondents in better health report lower desire (−0.063, 0.029). The absence of a positive income gradient in the share equation suggests that the health orientation of food expenditure in this sample is not simply a matter of budget.

**Endogeneity and robustness analysis**

Because consumers who prefer healthy food may be more likely to monitor themselves through quantitative means, reverse causality is a concern. We therefore instrument quantified-self engagement. Column (1) of Table 3 reports the instrumental-variables estimate, and the first-stage diagnostics indicate that the excluded instrument is relevant: the Kleibergen–Paap LM statistic is 19.801 and the Cragg–Donald Wald F statistic is 19.591, above the conventional threshold of 10.

Using the average quantified-self score of all other individuals of the same age across cities as an instrument for quantified self, the instrumental-variables estimate for the expenditure share is 0.116 (0.043), positive and statistically significant at the 1% level but smaller than the baseline ordinary-least-squares estimate of 0.128; because the instrument varies only across age cells, it captures common age-group variation and should be read as a check on the sign and persistence of the association rather than as an exact causal magnitude. We also conduct a placebo test, randomly permuting the quantified-self measure 1,000 times and re-estimating the model to obtain the distribution of coefficient estimates (Fig. 1). The baseline estimate lies well outside the placebo distribution, which is consistent with limited bias from omitted variables.

We also examine how the instrumental-variables estimate responds to alternative assumptions about the error structure, and report the corresponding statistics in the note to Table 3. The estimate remains positive and statistically significant at the 1% level under heteroskedasticity-robust standard errors, the convention used throughout the tables, and at the 5% level when standard errors are clustered on age cohort. City clustering is the least favourable case, because the sample contains only 12 clusters. These checks indicate that the estimate is driven neither by weak identification nor by the assumed error structure, although the city-clustered result should be read with caution.

**Table 3.** Endogeneity tests and robustness checks

| | IV | Quantile Estimation | | | Cluster-robust standard errors | EFA of quantified self | Health food desire | Eliminate expenditure of tobacco and alcohol |
|---|---|---|---|---|---|---|---|---|
| | (1) share | (2) Share(p25) | (3) Share(p50) | (4) Share(p75) | (5) share | (6) share | (7) desire | (8) share |
| selfqua | 0.116*** | 0.132*** | 0.130*** | 0.127*** | 0.128*** | 0.067*** | 1.484*** | 0.144*** |
| | (0.043) | (0.006) | (0.005) | (0.005) | (0.004) | (0.002) | (0.049) | (0.005) |
| constant | 0.254*** | 0.212*** | 0.222*** | 0.310*** | 0.254*** | 0.253*** | 0.190 | 0.298*** |
| | (0.027) | (0.039) | (0.035) | (0.036) | (0.029) | (0.027) | (0.322) | (0.033) |
| Control variables | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Occupation fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| City fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 1454 | 1454 | 1454 | 1454 | 1454 | 1454 | 1454 | 1454 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| adj. $R^2$ | | | | | 0.416 | 0.416 | 0.394 | 0.376 |
| Pseudo $R^2$ | | 0.251 | 0.242 | 0.244 | | | | |
| F statistic | | | | | | 31.49 | 28.73 | 26.77 |
| LM statistic | 19.801 | | | | | | | |
| CD Wald F statistic | 19.591 | | | | | | | |
| Sargan statistic | n.a. | | | | | | | |

Notes: Column (1) is exactly identified—one excluded instrument for one endogenous regressor—so the Sargan statistic has no degrees of freedom and is reported as n.a. The LM statistic and the CD Wald F statistic test underidentification and weak identification, respectively. The Anderson–Rubin 95% confidence interval is [0.043, 0.188], which excludes zero and is robust to weak identification. Alternative error structures give standard errors of 0.054 when clustering on age cohort ($p = 0.031$) and 0.070 when clustering on city ($p = 0.097$); with only 12 city clusters, wild cluster bootstrap-t p-values with the null imposed (Cameron et al. 2008) are 0.038 and 0.119, respectively.

We conduct additional robustness checks using quantile regression, city-level cluster-robust standard errors, an alternative measurement of quantified self, an alternative dependent variable, and a redefinition of the outcome denominator (Table 3). Columns (2) to (4) report quantile regression estimates at the 25th, 50th, and 75th percentiles (0.132, 0.130, 0.127, all significant at the 1% level), so the association is not driven by a particular part of the distribution. Column (5) clusters standard errors at the city level (0.128, 0.004) and Column (6) uses an exploratory factor analysis score (0.067, 0.002); both remain positive and significant. Column (7) uses stated desire as the dependent variable (1.484, 0.048) and Column (8) excludes tobacco and alcohol from the denominator (0.144, 0.005).

**Motivations and the quantified self–healthy food consumption relationship**

Using interaction terms, we examine whether quantified self motivations moderate the relationship between quantified self and healthy food consumption. The results are reported in Table 4.

The results show that self-optimisation motivation (opti) and its interaction with quantified self (selfqua × opti) are both positively and statistically significantly associated with the healthy-food expenditure share, suggesting that self-optimisation motivation strengthens the association between quantified self and healthy food consumption. A natural interpretation is that consumers who track with an improvement goal (for example, weight loss or chronic disease prevention) treat the resulting data as diagnostic feedback and adjust food choices in response, consistent with prior work emphasising goal pursuit and self-optimisation motives (Fan et al. 2022; Li and Zhang 2018). Quantified information about diet and health indicators may therefore be used more actively in purchase decisions (Jain et al. 2025).

Self-supervision motivation also strengthens the relationship between quantified self and healthy food consumption: in Table 4, both self-supervision motivation and its interaction with

quantified self (selfqua × supvi) are positive and statistically significant. One interpretation is that, when quantified self is used for self-monitoring and self-control (for example, tracking blood glucose or limiting sugar intake), individuals may place greater weight on quantified information about their health status when making food choices.

Self-enjoyment motivation is positively and statistically significantly associated with the healthy-food expenditure share, but its interaction with quantified self (selfqua × enjoy) is not statistically significant. Self-enjoyment motivation therefore shifts the level of healthy-food expenditure without conditioning the association. One interpretation is that, when quantified self is undertaken mainly for interest or enjoyment, it may increase attention in the short run but translate less directly into expenditure shares, particularly if the activity is treated more as recording than as goal-directed adjustment (Fan et al. 2022; Li and Zhang 2018). This is consistent with the view that enjoyment-oriented participation may not sustain the feedback seeking emphasised by self-optimisation and self-supervision motives (Swan 2009).

Overall, the interaction results indicate that self-optimisation and self-supervision motivations strengthen the relationship between quantified self and healthy food consumption, whereas self-enjoyment motivation does not; Hypotheses 3a and 3b are supported, while Hypothesis 3c is not. The coefficients on selfqua × opti and selfqua × supvi are almost identical in magnitude, which is consistent with Hypothesis 4.

**Table 4.** Motivations and the quantified self–healthy food consumption relationship

| | Healthy food expenditure share | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| selfqua | 0.118*** | 0.057*** | 0.122*** | 0.058*** | 0.123*** | 0.127*** |
| | (0.004) | (0.010) | (0.004) | (0.011) | (0.004) | (0.011) |
| opti | 0.023*** | 0.021*** | | | | |
| | (0.002) | (0.002) | | | | |
| selfqua* | | 0.024*** | | | | |
| opti | | (0.004) | | | | |
| supvi | | | 0.021*** | 0.018*** | | |
| | | | (0.002) | (0.002) | | |
| selfqua* | | | | 0.025*** | | |
| supvi | | | | (0.004) | | |
| enjoy | | | | | 0.015*** | 0.015*** |
| | | | | | (0.002) | (0.002) |
| selfqua* | | | | | | -0.002 |
| enjoy | | | | | | (0.004) |
| Constant | 0.214*** | 0.207*** | 0.205*** | 0.214*** | 0.217*** | 0.216*** |
| | (0.028) | (0.027) | (0.028) | (0.028) | (0.029) | (0.029) |

| | | | | | | |
|---|---|---|---|---|---|---|
| control variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Occupation fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| City fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 1454 | 1454 | 1454 | 1454 | 1454 | 1454 |
| adj. $R^2$ | 0.454 | 0.467 | 0.446 | 0.460 | 0.433 | 0.432 |
| F statistic | 35.45 | 36.36 | 34.40 | 35.37 | 32.64 | 31.72 |

Notes: Columns (1)–(2) use self-optimisation motivation, (3)–(4) self-supervision motivation and (5)–(6) self-enjoyment motivation; even-numbered columns add the interaction with quantified-self engagement. All columns include the control variables and occupation and city fixed effects (N = 1,454).

**Heterogeneity analysis**

We assess heterogeneity by interacting quantified-self engagement with each moderator and reporting the marginal effect at low and high moderator values, together with a test of whether the two marginal effects are equal. For continuous moderators the two values are one standard deviation below and above the sample mean; for binary indicators they are the two groups. Because the interaction coefficient equals the difference between the two marginal effects, the test of equality is the test that this coefficient is zero. Standard errors are computed by the delta method, and every specification retains the full control set, occupation fixed effects, and city fixed effects. We examine 22 moderators across four dimensions; the full results are reported in Supplementary Tables S1–S4 and summarised in Fig. 2.

Across the 22 moderators the marginal effect of quantified-self engagement on the healthy-food expenditure share is remarkably stable, staying between 0.118 and 0.134, and no difference between the low and the high group reaches statistical significance. Raw p-values range from 0.079 to 0.982 (median 0.729), and only the smoking contrast falls below 0.10 (0.118 versus 0.133, $p = 0.079$). Adjusting for multiplicity using the Holm (1979) and Benjamini–Hochberg (1995) procedures leaves no difference significant at either the 5% or the 10% level, the smallest adjusted p-value is 1.000 under Holm and 0.929 under Benjamini–Hochberg.

The absence of heterogeneity is itself informative, because the four dimensions we examine are precisely those along which the quantified-self mechanism is most likely to be conditioned. None of these margins conditions the association, suggesting that quantified-self engagement is linked to the health orientation of household food expenditure as a broad feature of the sample rather than through identifiable subgroups. This contrast is informative: motivations condition the relationship, whereas socio-demographic, technological, and dietary

characteristics do not. Three measurement caveats qualify the reading. The health-management-needs, food-safety-awareness, and food-nutrition-awareness columns rest on 1,079 rather than 1,454 observations because those items were not collected in all survey modules; duration and frequency of use are recorded on a four-category scale; and the technology, digital-ability, and awareness items are self-assessed, so measurement error attenuates their interaction terms and the corresponding null results are best read as upper bounds on heterogeneity.

## Conclusions

Using survey data on 1,454 respondents from 12 cities across 6 provinces in China, this paper examines the association between quantified-self engagement and healthy food consumption. After conditioning on a rich set of controls and occupation fixed effects—and in an instrumental-variables specification—quantified-self engagement remains positively and statistically significantly associated with both healthy food consumption desire and the healthy-food expenditure share. Self-optimisation and self-supervision motivations strengthen the association, whereas self-enjoyment motivation does not. Across heterogeneity specifications covering 22 moderator variables, the association is remarkably stable: the marginal effect on the healthy-food expenditure share ranges from 0.118 to 0.134 across low and high moderator values, and no group difference is statistically significant at the 5% level, whether raw or adjusted for multiple testing. The only estimate below the 10% threshold involves smoking status, for which the association is weaker among smokers (0.118 versus 0.133, $p = 0.079$); given the number of tests, this result is suggestive.

The paper links the emerging digital self-tracking literature to household food-demand economics through a money-metric outcome; it identifies the motivational conditions under which digital self-monitoring is most strongly associated with healthier food expenditure; and it documents that the association is a broad-based feature of the sample rather than one concentrated in identifiable demographic subgroups.

The first implication concerns how policy can support residents' participation in quantified self. Government agencies could improve the external environment, for example by encouraging hospitals and pharmacies through regulation or subsidies to offer free measurement of indicators such as blood glucose and blood lipids, which may help residents form monitoring habits. Related efforts could include supporting online health platforms alongside outreach that lowers barriers to participation. Public communication about diet and nutrition—such as

guidance on daily dietary requirements and food nutrient content—could further support residents' ability to interpret quantified information in food choices.

The second implication is that clarifying users' motivations matters for designing devices and apps. Self-optimisation and self-supervision motivations strengthen the relationship, while self-enjoyment motivation does not show a comparable moderating role. For providers, a practical priority is therefore to identify users' motivations, for example through onboarding assessments and observation of usage scenarios. Product design can then emphasise features supporting health-management goals, such as monitoring and recording health indicators, with enjoyment-oriented features—time-limited knowledge challenges, points and rewards, and community comparison—positioned as complements that sustain engagement.

The third implication concerns targeting. The heterogeneity results in Supplementary Tables S1–S4 provide no evidence that the association is concentrated among particular groups: the marginal effect of quantified-self engagement is essentially the same across age, gender, education, income, health status, household composition, technology-related attitudes and abilities, and dietary awareness. Policies that improve the external environment for quantified self and healthy consumption therefore need not target specific demographic segments; broad-based provision, such as the measurement services and nutrition-information platforms described above, is consistent with the pattern we observe.

Several limitations qualify these findings. First, the sample is a purposive quota sample rather than a probability sample, and campus- and park-recruited respondents may be more health-conscious than the general population. Second, the healthy-food expenditure share relies on respondent recall and on a definition of health-oriented food that partly depends on purchase motive, so measurement error in the share is possible. Third, the survey was administered both online and face-to-face, and mode effects cannot be fully ruled out. Fourth, the cross-sectional design and the strength of the identifying assumptions mean that the estimates should be interpreted as associations rather than causal effects.

## References

Almalki M, Gray K, Martin-Sanchez F (2016) Activity theory as a theoretical framework for health self-quantification: a systematic review of empirical studies. J Med Internet Res 18(5):e131

Anderson TW, Rubin H (1949) Estimation of the parameters of a single equation in a complete system of stochastic equations. Ann Math Stat 20(1):46–63

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57(1):289–300

Biroli P, Boneva T, Raja A, Rauh C (2022) Parental beliefs about returns to child health investments. J Econom 231(1):33–57. https://doi.org/10.1016/j.jeconom.2020.03.018

Cameron AC, Gelbach JB, Miller DL (2008) Bootstrap-based improvements for inference with clustered errors. Rev Econ Stat 90(3):414–427

Castro Campos B, Qi X (2024) A literature review on the drivers and barriers of organic food consumption in China. Agric Food Econ 12(1):18. https://doi.org/10.1186/s40100-024-00312-y

Cheng S, Lu Y, Ren Y, Jiang Z, Zhao M (2025) Which food date label brings us the most excellent opportunity to reduce food waste? Evidence from a quasi-natural intervention experiment in urban China. Agric Food Econ

Cragg JG, Donald SG (1993) Testing identifiability and specification in instrumental variable models. Econom Theory 9(2):222–240

Etkin J (2016) The hidden cost of personal quantification. J Consum Res 42(6):967–984. https://doi.org/10.1093/jcr/ucv095

Fan X, Fan J, Li J (2022) The effect of presentation characteristics of 'quantified self' data on consumers' continuance participation intention: an empirical study based on health-related apps. Psychol Res Behav Manag 15:2859–2877. https://doi.org/10.2147/PRBM.S381705

Firoozzare A, Boccia F, Yousefian N, Ghazanfari S, Pakook S (2024) Understanding the role of awareness and trust in consumer purchase decisions for healthy food and products. Food Qual Prefer 121

Gao Y, Zhang Z, Hu Y (2026) Digital bridges to nutritional transition: how internet penetration reshapes rural diets in China. Rev Dev Econ. https://doi.org/10.1111/rode.70195

Guo L (2016) Quantified-self 2.0: using context-aware services for promoting gradual behaviour change. University of Central Lancashire, Preston

Gupta A, Dhiman N, Yousaf A, Arora N (2020) Social comparison and continuance intention of smart fitness wearables: an extended expectation confirmation theory perspective. Behav Inf Technol

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6(2):65–70

Hou J, Huo X, Yin R (2019) Does computer usage change farmers' production and consumption? Evidence from China. China Agric Econ Rev 11(2):387–410. https://doi.org/10.1108/CAER-09-2016-0149

Jain A, Pradhan D, Kuanr A, Singh S (2025) Self-quantification and consumer well-being: a meta-analytic review. Psychol Mark 42:514–538. https://doi.org/10.1002/mar.22141

Kampen J (2020) From quantified self to qualified self: reducing academic procrastination through the qualified self. Master's thesis, University of Twente

Kersten-van Dijk ET, Westerink JHDM, Beute F, IJsselsteijn WA (2017) Personal informatics, self-insight, and behavior change: a critical review of current literature. Hum Comput Interact 32(5–6):268–296. https://doi.org/10.1080/07370024.2016.1276456

Kleibergen F, Paap R (2006) Generalized reduced rank tests using the singular value decomposition. J Econom 133(1):97–126

Kuhlgatz CH, Huang J, Antonides G (2023) Food demand and the nutrient intake of households in underdeveloped rural regions of China: an instrumental variable approach. China Agric Econ Rev 15(1):17–45. https://doi.org/10.1108/CAER-02-2022-0039

Lei L, Zhu Y, Liu Q (2021) Analysis on quantified self-behavior of customers in food consumption under the perspective of social networks. Complexity 2021:1–14. https://doi.org/10.1155/2021/6001654

Li D, Zhang Y (2018) Quantified self in the field of consumption: a literature review and prospects. Foreign Economics & Management 40(1):3–17. https://doi.org/10.16538/j.cnki.fem.2018.01.001

Lin HY, Hsu MH (2015) Using social cognitive theory to investigate green consumer behavior. Bus Strategy Environ 24(5):326–343. https://doi.org/10.1002/bse.1820

Liu R, Pieniak Z, Verbeke W (2013) Consumers' attitudes and behaviour towards safe food in China: a review. Food Control 33(1):93–104. https://doi.org/10.1016/j.foodcont.2013.01.051

Robertsson L (2014) Quantified self: an overview and the development of a universal tracking application. Dissertation, Umeå University

Rooksby J, Rost M, Morrison A (2014) Personal tracking as lived informatics. In: Proceedings of the 32nd annual ACM conference on human factors in computing systems. ACM, Toronto, pp 1163–1172

Ruckenstein M, Pantzar M (2017) Beyond the quantified self: thematic exploration of a dataistic paradigm. New Media Soc 19(3):401–418. https://doi.org/10.1177/1461444815609081

Sargan JD (1958) The estimation of economic relationships using instrumental variables. Econometrica 26(3):393–415. https://doi.org/10.2307/1907619

Schroeder TC, Tonsor GT, Pennings JME, Mintert J (2007) Consumer food safety risk perceptions and attitudes: impacts on beef consumption across countries. B E J Econ Anal Policy 7(1):1–27. https://doi.org/10.2202/1935-1682.1848

Shen X, Li Y (2017) An empirical investigation of factors affecting smart health device users' intermittent discontinuance. Journal of Management Science 30(1):31–42

Shin DH, Biocca F (2017) Health experience model of personal informatics: the case of a quantified self. Comput Hum Behav 69:62–74. https://doi.org/10.1016/j.chb.2016.12.019

Swan M (2009) Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine, and quantified self-tracking. Int J Environ Res Public Health

Yang Z, Paudel KP, Wen X, Sun S, Wang Y (2020) Food safety risk information-seeking intention of WeChat users in China. Int J Environ Res Public Health 17(7):2376

Ye J, Li H, Wu B (2011) Income inequality, consumption, and social-status seeking. J Comp Econ 39(2):191–204. https://doi.org/10.1016/j.jce.2010.12.004

Zhang Y, Li D, Zhang C, Zhang H (2019) Quantified or nonquantified: how quantification affects consumers' motivation in goal pursuit. J Consum Behav 18(2):120–134. https://doi.org/10.1002/cb.1752

Zhang Y, Zhang H, Zhang C, Li D (2020) The impact of self-quantification on consumers' participation in green consumption activities and behavioral decision-making. Sustainability 12(10):4098. https://doi.org/10.3390/su12104098

# Figure

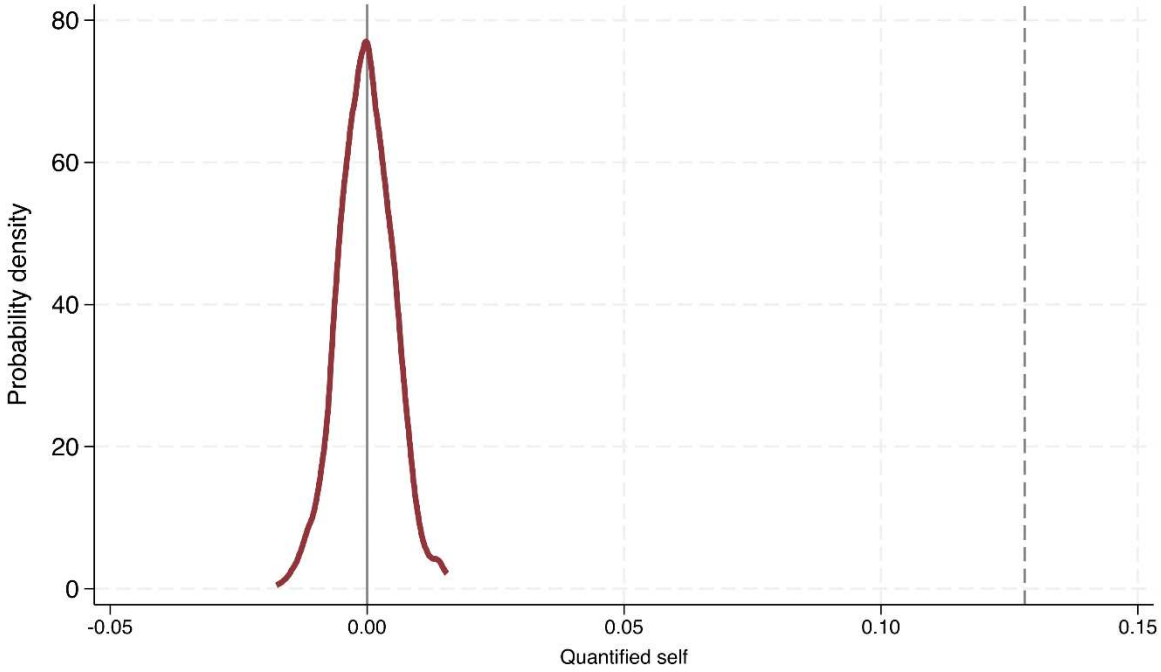



**Figure 1.** Distribution of permuted coefficients in the placebo test.

Notes: The figure shows the probability density of the coefficient on quantified self estimated from 1,000 regressions in which the quantified-self measure is randomly permuted. The solid vertical line marks zero and the dashed vertical line marks the baseline estimate of 0.128.

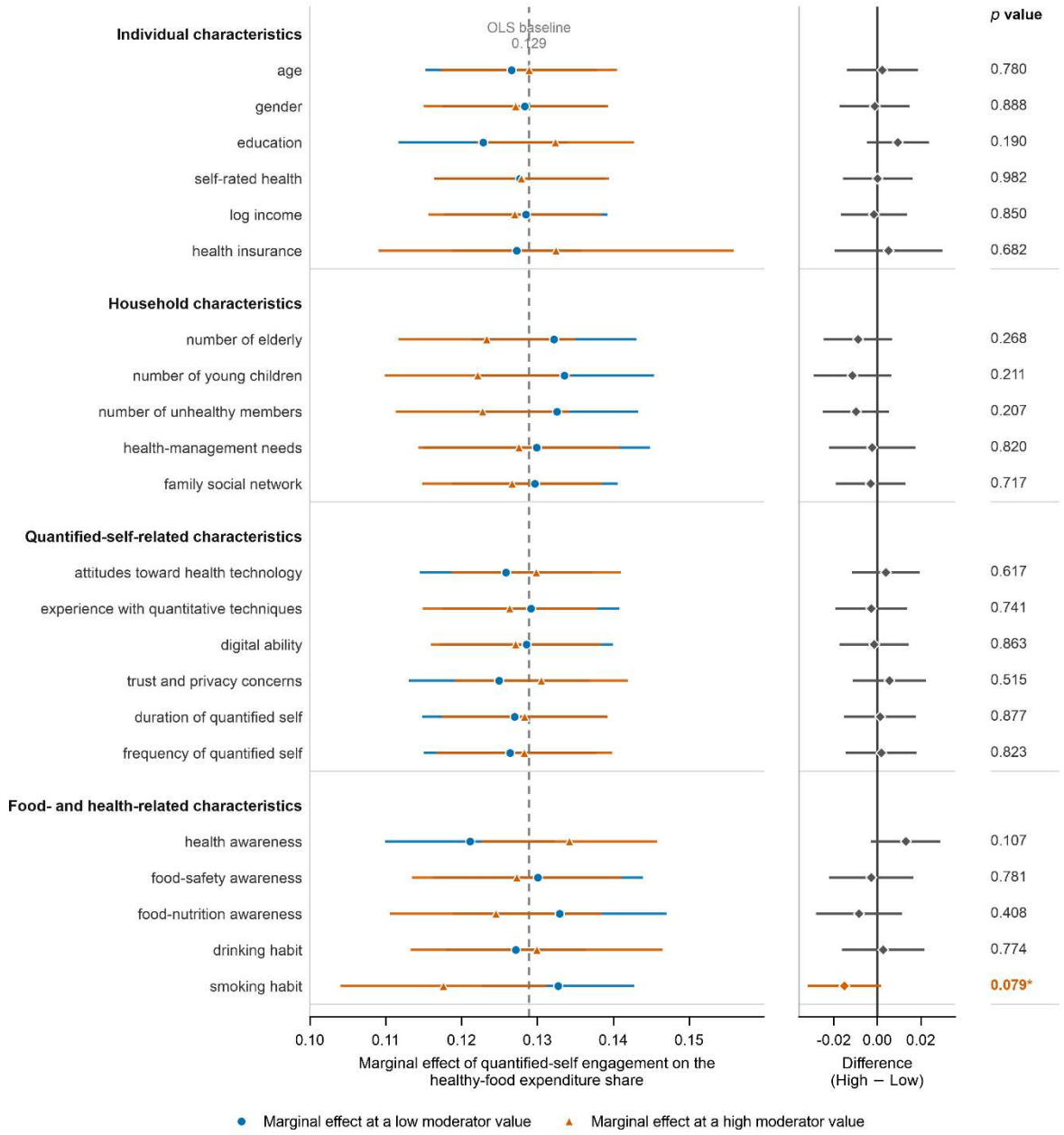



**Figure 2.** Marginal effects of quantified-self engagement at low and high values of each of the 22 moderators.

Notes: Circles and triangles show the marginal effect at a low and at a high moderator value with 95% confidence intervals, and the grey dashed line marks the baseline estimate. Diamonds in the right-hand panel show the difference between the two marginal effects, and the asterisk marks the smoking contrast, the only difference with $p < 0.10$.

## Data availability

All data and materials underpinning this study are provided with the manuscript as supplementary files. No material is withheld, and no access restriction, permission or fee applies to any of them.

Specifically, the data package contains: (i) **the data, in both raw and processed form** — the raw respondent-level file as compiled at the end of fieldwork (1,454 respondents × 113

variables) and the analysis-ready respondent-level file used for every estimate reported here (1,454 respondents × 99 variables); the processed file differs from the raw file only in that 14 redundant estimation-sample mask columns were dropped, with no respondent removed, no value imputed and no variable recoded; (ii) **the survey instrument and fieldwork protocol** — the questionnaire as administered and the sampling and fieldwork procedures, also reproduced as Supplementary Note 1; (iii) **variable details** — a variable dictionary that lists, for every variable, its original label, its English description, its role in the paper, its type, its missing rate, its univariate statistics and its admissible values; (iv) **the analysis resources** — the complete analysis code (Stata do-files and Python scripts), the order in which the scripts are run, the required Stata packages and Python package versions, and the intermediate outputs from which Tables 1–4, Supplementary Tables S1–S5 and Fig. 2 are generated; and (v) **supporting documentation** — a README documenting every processing step taken and a cell-by-cell reproducibility check of all numbers printed in this manuscript against the code output (329 cells compared: 327 exact, 2 differing only in the second decimal of an F statistic, and no mismatch).

The released data are fully de-identified. Names and telephone numbers were recorded during fieldwork solely for quality control and follow-up verification, were stored separately from the questionnaire responses, and were deleted before the analytic dataset was constructed. Neither released file contains a name, telephone number, address or any other identifying field, and no individual participant is identifiable in this manuscript or in any supplementary file. Because no material needs to be withheld, there are no data or materials underlying this study that cannot be made available.

## Competing interests

The authors declare no competing interests.

## Ethics declarations

**Ethical approval.** The authors confirm that all research was performed in accordance with relevant named guidelines and Helsinki regulations, when human participants are involved. Ethics approval was obtained from the Professor Committee at the International Business School of Shaanxi Normal University, China. The ethical approval protocol number 2025D035.

Explicit and informed consent was obtained from all participants whose data was collected. They were also informed that all private information, including their names and companies' names, would be anonymized in the study results.

**Informed consent.** Informed consent was obtained from every participant before data collection. Interviewers read a standard instruction script aloud, explained that participation was voluntary and could be withdrawn at any time without consequence, and obtained a signed consent form from each respondent. Only adults aged 18 or over were included in the study.

**Identifying information.** Names and telephone numbers were recorded during fieldwork solely for quality control and follow-up verification, were stored separately from the questionnaire responses, and were deleted before the analytic dataset was constructed. The dataset analysed and released contains no identifying information, and no participant is identifiable in this manuscript or in the supplementary files.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 07SupplementaryAppendixHSSCfinal.docx