The Consistency of Rationality Measures

Mingshi Chen, Tracy Xiao Liu, You Shan, Songfa Zhong, and Yanju Zhou*

February 23, 2023

Abstract

Revealed preference analysis provides a powerful method for measuring rationality—the extent to which choice data of individual can be rationalized by some well-behaved utility function. Whereas rationality has been widely measured in various lab and field settings, the consistency of these measures has not been examined. We combine budgetary decisions in the lab and food decisions in the field based on scanner data to measure the rationality of individual consumers in a large grocery store. We show that the rationality score for risky or social decisions in the lab is uncorrelated with that of food decisions in the field. By contrast, the rationality score is highly correlated between risky and social decisions in the lab, as well as between food decisions in the lab and in the field. We further show that behavioral factors including purchasing experience, personality traits and cognitive skills may underlie rationality scores across different environments.

Keyword: revealed preference, rationality, generalizability, experiment

JEL classification: C91, D81, D91

*Chen, Liu and Shan: School of Economcis and Management, Tsinghua University (Chen: cms19@mails.tsinghua.edu.cn; Shan: shany19@mails.tsinghua.edu.cn; Liu: tracy.umich@gmail.com). Zhong, National University of Singapore and New York University Abu Dhabi (zhongsongfa@gmail.com). Zhou, Central South University (zyj4258@sina.com). Liu gratefully acknowledges financial support by NSFC (72222005) and Tsinghua University (No. 2022Z04W01032). Zhou gratefully acknowledges financial support by NSFC (71991463).

1 Introduction

Central to economic analysis is the assumption that a decision maker (DM) maximizes her utility function given her budget constraint (Samuelson, 1938). Revealed preference analysis is a powerful method for characterizing the conditions under which DM with a given set of choice data indeed maximizes some well-behaved utility function (Afriat, 1967, 1972; Varian, 1982, 1990). Based on revealed preference analysis, the degree of consistency with utility maximization has been widely used as a natural criterion to measure rationality in various settings.[1] In experimental literature, it is increasingly common to compute rationality score based on risky, intertemporal, and social budgetary decisions before examining specific utility functions (see, for example, Ahn *et al.*, 2014; Andreoni and Miller, 2002; Choi *et al.*, 2007; Echenique *et al.*, forthcoming; Fisman *et al.*, 2007; Halevy *et al.*, 2018; Polisson *et al.*, 2020). Going beyond the lab, researchers also examine rationality score of consumers in the field settings using expenditure data from surveys and scanner data from grocery stores (see, for example, Blundell *et al.*, 2003, 2008; Dean and Martin, 2016; Echenique *et al.*, 2011). It has been further suggested that rationality scores capture the ability to make good decisions and have been proposed as an important determinant of wealth differences across individuals in a general population (Choi *et al.*, 2014) and development gaps between countries (Cappelen *et al.*, 2021). Although these studies provide important insights into the notion of rationality and its applicability in various settings, many questions remain unresolved.

One important question is whether individuals with higher rationality scores as subjects in the lab will also have higher rationality scores as consumers in the field. A perhaps narrower and related question is whether individuals as subjects in the lab will have consistent rationality scores across various settings such as risky, social, and consumption goods deci-

[1] In the literature, rationality is often broadly defined to encompass alternative decision rules and heuristics (Kahneman, 2003; Simon, 1979; Thaler, 2016). For simplicity, we follow some prior studies in the revealed preference analysis and use "rationality" to refer to the extent of consistency with utility maximization. See, for example, Chambers and Echenique (2016), Crawford and De Rock (2014), and Echenique (2020) for reviews of recent developments in the revealed preference literature.

sions. On the one hand, it is conceivable that rationality scores can be generalized across settings, because it may capture the general capacity to make good decisions. On the other hand, it is also likely that rationality scores are specific to a given setting, because the same individual in different settings may have different preference structures, face different budget constraints, and have different experiences in making decisions, which can give rise to different rationality scores inferred from choice data. In this paper, we address these questions by combining laboratory experiments and scanner data and examining the consistency of individual consumers' rationality measures in a large grocery store.

In the first study (Study 1), we examine the rationality of the same individuals making shopping decisions in the (grocery store) field and making risky decisions in the lab. To compute rationality scores in the field, we restrict our attention to 6,126 customers of whom we have purchase records of meat and vegetables—the two most common consumption categories—over 24 consecutive months, so that we have sufficient power for the revealed preference analysis (for further discussion, see Section 2). To measure rationality in the lab, we successfully invited 1,073 of the 6,126 customers in the field sample to participate in a budgetary experiment in which subjects make 22 portfolio decisions between two Arrow securities with different budget lines (Choi *et al.*, 2007, 2014; Halevy *et al.*, 2018; Kim *et al.*, 2018). In the experiment, subjects allocate 100 experimental points between two accounts in which points are converted to cash with different exchange rates, and the amount in each account will be paid out with 50 percent chance. Our experimental design enables us to compute individual rationality scores in the field based on purchase records for meat and vegetables over 24 consecutive months and in the lab based on 22 portfolio decisions between two Arrow securities.

We measure rationality by evaluating the consistency of individual choices with the Generalized Axiom of Revealed Preference (GARP), a necessary and sufficient condition whereby a data set can be rationalized by a well-behaved utility function (in accordance with utility maximization). To assess how closely an individual choice data set complies with GARP, in addition to the commonly used critical cost efficiency index (CCEI, Afriat, 1972), several measures have been proposed in the literature, including the Houtman-Maks index (HMI, Houtman and Maks, 1985); money pump index (MPI, Echenique *et al.*, 2011); and minimum cost index (MCI, Dean and Martin, 2016).[2] We compute individual rationality scores using each of these four indices.

Our primary finding is that the rationality scores are uncorrelated between food decisions in the field and risky decisions in the lab. This observation is robust to the use of specific indices with respect to the extent of GARP violations, namely, CCEI, HMI, MPI, and MCI, since the Spearman correlation ranges from 0.1 percent to 4 percent. Not surprisingly, these four indices are highly correlated within the field (with the Spearman correlation ranging from 56 percent to 90 percent), as well as within the lab (75 percent to 98 percent). Whereas there have been theoretical discussions on how best to capture the extent of GARP violations (Echenique, 2021; Polisson and Quah, 2022), here we show empirically that these different measures are highly correlated.

We then conduct several robustness checks for the observed low correlation between field and lab. For rationality scores in the field, we examine several alternative measures, such as adding fruits as a third consumption category in addition to meat and vegetables; extending the analysis to 36 months instead of 24 months; using alternative methods to construct the budget lines (Dean and Martin, 2016; Echenique *et al.*, 2011); adjusting scores based on the power of individual budget lines (Selten, 1991); applying the revealed price preference (Deb *et al.*, 2022); and using downward-sloping demand to proxy individual rationality (Echenique *et al.*, forthcoming). For rationality scores in the lab, we also check several alternative measures by imposing more restrictions on the utility functions such as respecting first-order stochastic dominance or expected utility (Nishimura *et al.*, 2017; Polisson *et al.*, 2020). Regardless of these alternative measures, the correlation between rationality scores in the

[2] We provide a brief review of revealed preference analysis including GARP and each of these four indices in Section 2.

field and in the lab continues to be close to zero.

We further examine the demographic and behavioral correlates for the rationality scores. While we find that rationality scores in both the lab and the field are positively correlated with family income and negatively correlated with age, they differ in several correlates. Rationality scores in the field are positively correlated with family size, the amount of expenditure, and purchasing frequency, as well as conscientiousness scores of the Big Five personality traits; rationality scores in the lab are positively correlated with IQ scores using seven questions from Raven's Progressive Matrices and negatively correlated with the neuroticism scores of the Big Five personality traits. These results suggest that certain behavioral correlates may underpin the difference between rationality scores in the two settings.

Study 1 shows that the rationality scores for risky decisions in the lab and for food decisions in the field exhibit no correlation, and the question remains whether the low correlation is due to difference in the types of decisions—risky decisions versus food decisions—or due to the difference in lab and field settings. To answer this question, we conduct a second experiment (Study 2). We invite another 305 customers to participate in three budgetary experiments. In addition to the 22 risky decisions between two Arrow securities used in Study 1, subjects make 22 social decisions on allocating experimental points between herself and an anonymously paired subject (Andreoni and Miller, 2002; Fisman *et al.*, 2007) and 22 food decisions on allocating money between one type of vegetable, tomatoes, and one type of meat, ham. These three budgetary experiments enable us to measure rationality scores for risky, social, and food decisions in the lab. Moreover, we can also measure the rationality of these subjects based on our available scanner data.

We summarize Spearman's correlation coefficients of CCEI scores in these four settings in Table 1 below. First, we replicate the observed low correlation of rationality scores between food decisions in the field and risky decisions in the lab in Study 1 and generalize the result by showing that correlation of rationality scores between food decisions in the field and social decisions in the lab is also low. Second, by contrast, we find that rationality scores are highly correlated between risky decisions and social decisions in the lab as well as between food decisions in the lab and in the field. Moreover, the correlation is moderate between food decisions in the lab and risky/social decisions in the lab. Similar to Study 1, we show that these correlational patterns are robust with respect to alternative rationality measures, and some behavioral correlates may underpin the difference in rationality scores across settings.

Table 1: Correlations of the Rationality between Domains

	Risk		Social	Food: lab
	Study1	Study2	Study2	Study2
Food: field	0.001	0.090	0.061	0.580
	(0.978)	(0.239)	(0.429)	(0.000)
Risk			0.396	0.237
			(0.000)	(0.000)
Social				0.237
				(0.000)

Note: Spearman's ρ are reported. P-value in parentheses.

Our study adds to the literature on measuring rationality based on revealed preference analysis. In addition to theoretical and experimental studies, rationality indices have been proposed as a measure of decision-making quality and widely used in applied settings, which link individual rationality to education, occupation, income, and wealth (e.g., Banks *et al.*, 2019; Cappelen *et al.*, 2021; Carvalho *et al.*, 2016, 2019; Choi *et al.*, 2014; Kim *et al.*, 2018; Li *et al.*, 2023). In this literature, most studies focus on a specific type of decisions either in the lab or in the field. One exception is Kim *et al.* (2018)'s study, in which they collect choice data on both risky and intertemporal decisions in the lab to evaluate the effects of an education program. Based on the rationality scores of risky and intertemporal decisions in their data set, we compute the Spearman's correlation to be 51.4 percent, which is consistent with the observed correlation between risky and social decisions in our experiment. Our study is the first to examine the consistency of rationality measures, and sheds light on understanding of rationality and its applications.

Our study also contributes to the literature on the consistency of choice behavior across settings. A question that has attracted great attention is the consistency of risk preferences. For example, Weber *et al.* (2002) shows that risk attitudes are inconsistent across domains; these include financial, health/safety, recreational, ethical, and social decisions. Dohmen *et al.* (2011) finds that different measures of risk attitudes are imperfectly correlated across contexts. Barseghyan *et al.* (2011) and Barseghyan *et al.* (2013) use three deductible choices made in the domain of auto and homeowner insurance to estimate individual risk attitude in each domain using structural approach and reject the null of fully domain-general risk attitudes. Einav *et al.* (2012) provides evidence in support of consistency within five employer-provider insurance coverage decisions, and inconsistency between these insurance decisions and investment decisions with respect to a 401(k) plan. In addition to risk preferences, moderate or weak consistency across settings has also been reported in time preferences (Augenblick *et al.*, 2015; Burks *et al.*, 2012); social preferences (Cohn and Maréchal, 2018; Dai *et al.*, 2018; Galizzi and Navarro-Martinez, 2019); and strategic sophistication (Georganas *et al.*, 2015; Levitt *et al.*, 2011; Rubinstein, 2016). Moreover, our study adds to a growing literature on the generalizability of observations in the lab to field settings (see., for example, Camerer, 2011; Levitt and List, 2007a,b, for summaries and discussions).

Building on these studies, here we address an important yet unexplored question—the consistency of rationality measures—and examine this question in two experiments across risky, social, and food decisions in the lab and in the field. On the one hand, the lack of correlation between lab and field suggests that rationality scores revealed from choice behavior in various environments may not solely reflect the general ability to make good decisions according to one's preferences. On the other hand, we do observe high correlations between risky and social decisions in the lab, and between food decisions in the lab and in the field, which indicates that rationality measure is not fully context-dependent whereby each choice context has its own rationality measure. In this regard, This finding supports the idea that the ability to make good decisions is multidimensional (Currie and MacLeod, 2017; Deming, 2021), and that choice behaviors are interconnected and can be reduced to several underlying common factors (Chapman *et al.*, 2023; Dean and Ortoleva, 2019; Stango and Zinman, 2022).

The rest of the paper is organized as follows. Section 2 presents the theoretical background of revealed preference analysis, Sections 3 and 4 present the experimental design and results of the two studies, and we discuss our findings in Section 5.

2 Theoretical Background

Consider a decision maker who chooses bundles $x^i \in \mathbb{R}^K_+$ from budget lines $\{x : p^i \cdot x \leq p^i \cdot x^i, p^i \in \mathbb{R}^k_{++}\}$. A *data set* $\mathcal{O} = \{(p^i, x^i)\}_{i=1}^N$ refers to a collection of the N decisions of the decision maker. Let $\mathcal{X} = \{x^i\}_{i=1}^N$ be the set of bundles chosen by the decision maker. We say that x^i is *directly revealed preferred* to x^j, denoted by $x^i \succsim^* x^j$, if a subject chooses x^i when $x^j \in \mathcal{X}$ is affordable (i.e., $p^i \cdot x^j \leq p^i \cdot x^i$). Denote the asymmetric part as $\succ^*$, which refers to the relation of *directly strictly revealed preference.* Denote $\succsim^{**}$ the transitive closure of $\succsim^*$, which refers to the *revealed preferred* relation. A data set $\mathcal{O}$ satisfies the Generalized Axiom of Revealed Preference (GARP) if the following holds:

$$\text{for all } x^i \text{ and } x^j,\ x^i \succsim^{**} x^j \text{ implies } x^j \not\succ^* x^i. \tag{1}$$

We say that a utility function $U : \mathbb{R}^k_+ \to \mathbb{R}$ *rationalizes* the data set $\mathcal{O}$ if for every bundle x^i :

$$U(x^i) \geq U(x) \text{ for all } x \in \mathbb{R}^K_+ \text{ s.t. } p^i \cdot x \leq p^i \cdot x^i.$$

Afriat's theorem (Afriat, 1967; Varian, 1982) states that a data set can be rationalized by some well-behaved (continuous and strictly increasing) utility function, if and only if the data set obeys GARP. Afriat's theorem is a powerful tool for analyzing choice behavior.

2.1 Indices of Rationality

Afriat's theorem informs whether a data set can be rationalized by some utility function, but it provides no information on the degree of rationalizability. Several indices have been developed to allow us to quantify the extent of such violations. Below we briefly review four of them.

Critical Cost Efficiency Index (CCEI): A popular approach for measuring the departure from rationality is the *critical cost efficiency index* (CCEI) proposed by Afriat (1972). A subject has a CCEI $e \in [0, 1]$ if e is the largest number with a well-behaved U that rationalizes the data set for every $x^i \in \mathcal{X}$:

$$U(x^i) \geq U(x) \text{ for all } x \in \mathbb{R}^K_+ \text{ s.t. } p^i \cdot x \leq ep^i \cdot x^i. \tag{2}$$

A CCEI of 1 indicates passing GARP perfectly. A CCEI less than 1—say, 0.95—indicates that there is a utility function for which the chosen bundle x^i is preferred to any bundle that is more than 5 percent cheaper than x^i. Put differently, the CCEI can be viewed as the amount by which a budget constraint must be relaxed in order to remove all violations of GARP, because the decision maker can achieve her utility targets by spending less money (Afriat, 1972; Varian, 1990).

Houtman-Maks Index (HMI): Houtman and Maks (1985) proposes an alternative approach by measuring the maximal number of observations in the observed sample that are consistent with rational choice. For example, an HMI score of 0.941 indicates that the largest proportion of subsets of choices consistent with GARP is 94.1 percent.

Money Pump Index (MPI): Echenique *et al.* (2011) provides an index to measure the amount of money one can extract from an individual for each violation of GARP. Their measure is based on the idea that an individual with a GARP violation can be exploited by an "arbitrager" as a "money pump". The "arbitrager" can choose allocation x^1 at price p^2 and allocation x^2 at price p^1, then trade x^1 with the individual at p^1 and x^2 at p^2, which yields a profit of $p^1(x^1 - x^2) + p^2(x^2 - x^1)$. Given multiple violations of GARP, a money pump cost will be associated with each violation. Following Echenique *et al.* (2011), we use the mean money pump costs for cyclic sequences of allocations with the length of two. For example, an MPI score equal to 0.059 means on average 5.9 percent of expenditure can be exploited by an "arbitrager" from GARP violation.

Minimum Cost Index (MCI): Dean and Martin (2016) proposes MCI to measure the minimum cost of breaking all revealed preference cycles in a data set. MCI is defined as $\min_{B \subset R_0} \frac{\sum_{(x,y)\in B} p_x(q_x - q_y)}{\sum_{x=1}^{T} p_x q_x}$ such that R_0/B is acyclic. The index has a high value when there are a large number of cycles for which all GARP violations are based on significant monetary differences relative to total expenditure. Thus, the MCI responds to both the number and severity of revealed preference violations. For example, if the MCI is 0.004, it means that the average cost of preference relations that must be removed to render the data set consistent with rationality is 0.4 percent.

To summarize, these indices measure the degree of GARP violations from different perspectives. Detailed theoretical discussions can be found in Echenique (2021) and Polisson and Quah (2022).

3 Study 1

In this section we first describe the field data and experiment. We examine the choice behavior of consumers who shop at a leading retail company with more than 400 supermarkets spread over 6 provinces in southern China. We obtained a user-level scanner panel data set from the company to measure rationality in food preference, and conducted an experiment to measure rationality in risk preference. The combination of scanner panel data and experimental data enable us to test the consistency of rationality measures.

3.1 Scanner Data Set

We have access to individual transaction data for consumers using their retail company membership to purchase from any of the company's supermarkets. For each observation, we have a unique membership ID, item index, transaction date, consumed units, the amount of expenditure, coupon value, store ID, and whether the transaction is online or in person.

We focus on two types of food consumption, meat and vegetables in the analysis for the following reasons. First, because these are the mostly frequently consumed goods at supermarkets, we are able to garner a sufficient number of observations. In particular, we restrict our attention to highly frequent customers who have purchase records for 24 consecutive months from January 2019 to December 2020. Second, since the unit for meat and vegetables is kilograms, it is convenient to have a refined measure of the quantity. We measure the price and quantity for meat and vegetables by averaging the quantity and price of the items in each category. We denote q_{itn} as the quantity of item i in consumption category I for consumer n in month t, and Q_{Itn} as the quantity of consumption category I for consumer n in month t, which is obtained by the sum of q_{itn}.

$$Q_{Itn} = \sum_{i \in I} q_{itn}.$$

In a similar vein, we denote P_{Itn} as the price of the consumption category I for consumer n in month t, which is obtained by averaging the prices weighted by quantity:

$$P_{Itn} = \frac{\sum_{i \in I} e_{itn}}{Q_{Itn}},$$

where e_{itn} is the expenditure for item i in category I. With quantity Q_{Itn} and price P_{Itn}, we can construct the budget lines for consumer n in month t and perform revealed preference analysis for each consumer n.[3]

It is worth noting that it is often challenging to construct individualized price and quantity using consumption data. For example, Echenique *et al.* (2011) assumes the same price for

[3]To perform the revealed preference analysis, we need to assume that preference over meat and vegetables is weakly separable from other goods and services (see for example, Dean and Martin (2016) and Echenique *et al.* (2011)).

each item for all consumers and Dean and Martin (2016) assumes the same price for all items under the same good category for all consumers. Since we are able to access high frequency data at the individual level, we can construct individualized indices for each customer. This helps to reduce measurement error and increases the power of GARP tests. We also use the methods of Dean and Martin (2016) and Echenique *et al.* (2011) as robustness checks.

3.2 Lab: Risk Preference

We measure the consumer's rationality of risk preference using a standard budgetary design. In each round of the task, participants allocate 100 tokens between two contingent assets (account) and receive the money in one asset with equal probability. The exchange rate between token and cash differs for the two assets and varies across rounds. If subjects are risk neural or risk seeking, they would allocate all the tokens to the cheaper asset; if they are risk averse, they would allocate some tokens to each asset depending on their risk preference.

Moreover, to help supermarket consumers understand the experimental tasks, we discretize the choice set as in Kim *et al.* (2018), whereby subjects choose between 11 allocation options.[4] Figure 1 presents a screenshot of the task. In total, participants face 22 decision tasks presented in random order, and their choices in one round are randomly selected for final payment. Different from Choi *et al.* (2007, 2014), we do not use randomly generated budget lines for each subject. Instead, every subject is presented with the same set of budget lines to ensure that the power for revealed preference tests is the same for all subjects, which facilitates comparison across subjects (Halevy *et al.*, 2018). At the end of the experiment, we include a 10-question version of Big Five personality questions, seven questions from Raven's progressive matrices, and collect demographic information. Detailed instructions for the experiment are provided in Appendix C.

[4]Polisson and Quah (2013) shows that Afriat's theorem is applicable in discrete settings, as long as the consumer obeys cost efficiency.

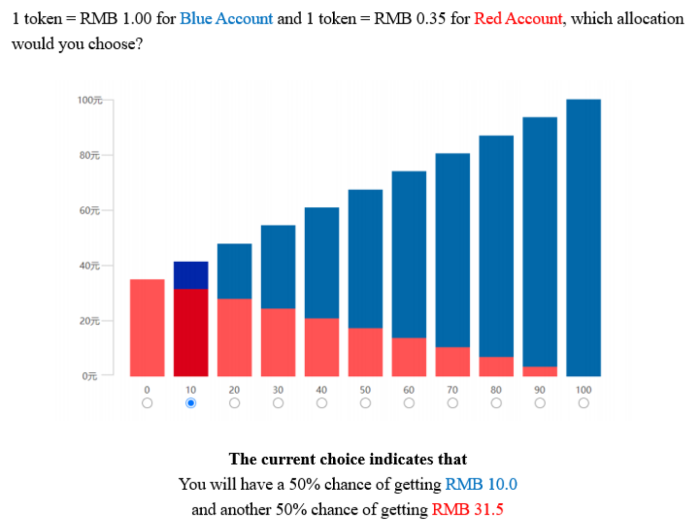


Figure 1: Screenshot for the Risk Preference Task

We conducted the experiment from July to October 2021. Of the 6,126 frequent consumers in our food consumption analyses, we have contact information for 3090. We invited these consumers to participate in an incentivized experiment in nearby supermarkets. In total, 1073 customers participated in our experiments across 96 supermarkets. All subjects received RMB 50 (USD 1 = RMB 6.69) for participating. Moreover, we randomly paid 10 percent of subjects based on their choices in the risk task. On average, each subject spent about 20 minutes to finish the experiment and received an average payment of RMB 57. The experiment was approved by the institutional review board of the National University of Singapore.

Table A1 in Appendix A reports summary statistics for 1,073 subjects for whom we have both scanner and lab data. Of these, 75.1 percent are female, 70.8 percent were born between 1970 and 1989, and more than 50 percent have a high school education or above. In terms of family size, the majority have at most 4 family members. Median family income is RMB 5000-10000 per month. The average monthly expenditure in this supermarket is RMB 485.6 and the average monthly purchase frequency is 11.8 days. Also, when comparing these subjects with the full sample of 6,126 consumers, we do not find significant differences in demographic variables, and we thus focus on these 1,073 subjects in later analyses.

3.3 Study 1: Results

3.3.1 Descriptive Statistics of the Rationality Measure

Figure 2 presents the cumulative distribution of rationality scores for food preference (solid line) and the summary statistics is reported in Table A2. We focus on the discussion for CCEI and the data pattern for other measures is similar. We find that 14.5 percent of customers have no violations of GARP in their choices for food consumption, and the average CCEI score is 0.941, which implies that the budget must be reduced by 5.9 percent to remove all GARP violations. To examine the power of the test, we apply Bronars' index and generate observations using uniformly random allocations along each of the budget lines. We compute the average CCEI and plot the cumlative distribution (dashed line). In contrast to the CCEI based on choice data, data from simulated random allocations have substantially lower rationality scores, for example, only 6.6 percent of their CCEI scores are over 0.9. This suggests that our choice data have sufficient power to detect GARP violations.[5]

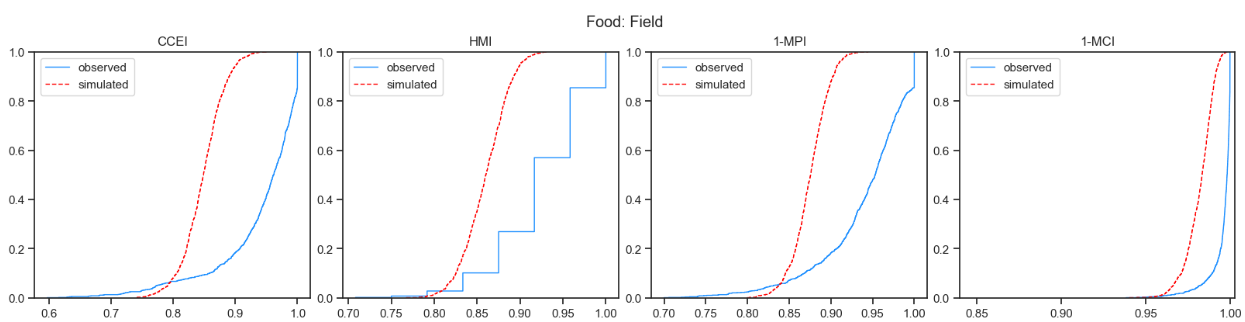


Figure 2: Distribution of the Rationality Index of Food Preference

[5]Although the observed CCEI in our field data is high, it is lower than those reported in some prior studies. For example, the average CCEI is 0.976 in Echenique *et al.* (2011), and 0.990 in Dean and Martin (2016). This could be due in part to the difference in the power of the test. For example, Echenique *et al.* (2011) shows that the number of GARP violations frequency is either 3 or 4 out of 494 households for the randomly generated data set, which tends to be lower than that observed in the actual data set.

Moreover, Figure 3 presents the cumulative distribution of rationality index (solid line) and Bronars' index (dashed line) for risk preference in the lab (see Table A3 for summary statistics). We find that 22.2 percent of subjects have no violations of GARP with an average CCEI score of 0.943. Comparison of distributions of the CCEI for actual choices and randomly generated choices is similar to that for food preference, which suggests that the chosen budget lines have sufficient power for revealed preference analysis.[6]

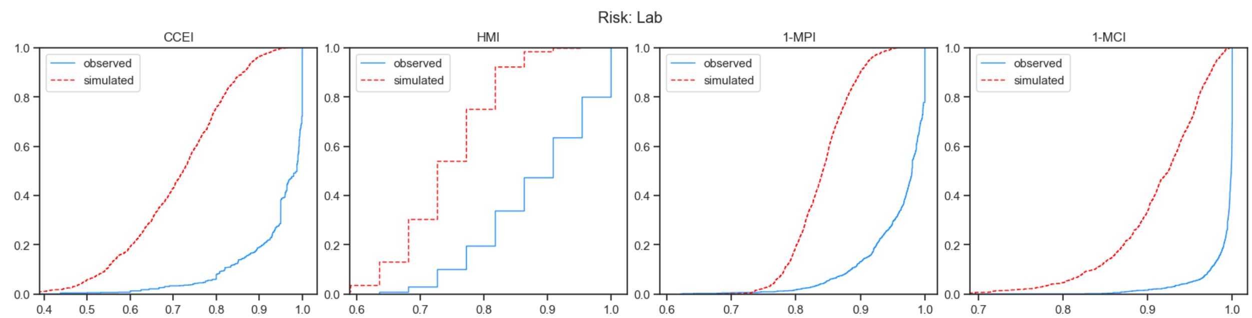


Figure 3: Distribution of Rationality Index of Risk Preference

In addition, we check the consistency of these four measures on the degree of GARP violation using Spearman's correlation test. Table 2 reports correlation coefficients for pairwise comparisons and shows that the correlation coefficients are 0.56 or higher within each type of preferences, which suggests high consistency between different measures. In subsequent analyses, for ease of presentation, we focus on results using CCEI and report results for other indices in the Appendix as robustness checks.

[6]The observed CCEI distribution is comparable to those in other experimental studies on risk preferences. For example, the average CCEI is 0.881 in Choi *et al.* (2014), 0.937 in Choi *et al.* (2007), and 0.979 in Halevy *et al.* (2018).

Table 2: Correlation of CCEI, HMI, MPI and MCI in Study 1

	Food				Risk		
	CCEI	HMI	MPI		CCEI	HMI	MPI
HMI	0.673			HMI	0.754		
	(0.000)				(0.000)		
MPI	-0.898	-0.563		MPI	-0.909	-0.636	
	(0.000)	(0.000)			(0.000)	(0.000)	
MCI	-0.902	-0.760	0.798	MCI	-0.983	-0.774	0.900
	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)

Note: P-values are in parentheses.

3.3.2 Correlations of Rationality: Food vs. Risk Preference

We now turn to the main research question regarding the consistency between the rationality of food preference in the supermarket and risk preference in the lab. Figure 4 presents a scatter plot of the two and shows that the correlation coefficient is 0.001 ($p > 0.1$). That is, consumers with a high level of rationality in the risk task experiment do not necessarily exhibit a high level of rationality in food preference.

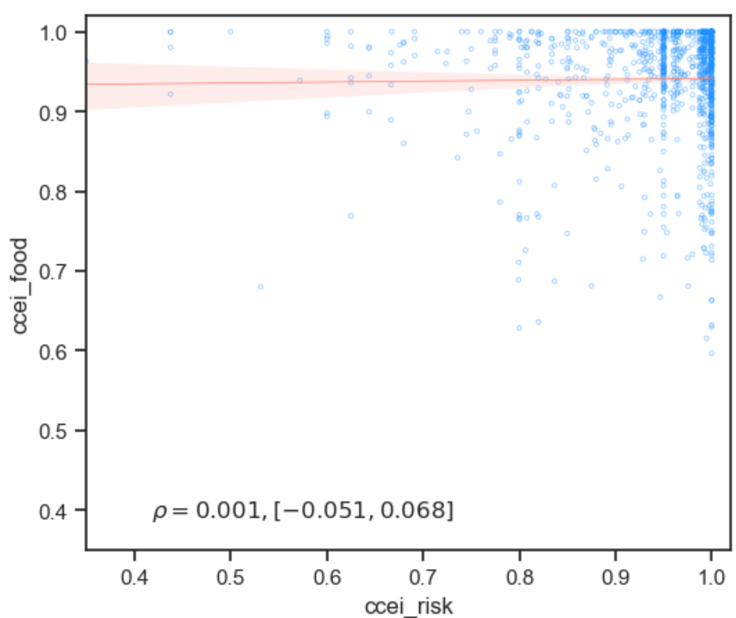


Figure 4: Correlations of CCEI in Study 1: Food vs. Risk Preferences

To further validate this low correlation, we conduct the following analyses. First, we consider alternative measures of the rationality for the scanner dataset, by (1) assuming the price is the same for all customers in each month (Dean and Martin, 2016; Echenique *et al.*, 2011);[7] (2) extending the time window from 2019-2020 to 2018-2020; (3) adding fruits as the third consumption category; (4) using Selten's score to adjust the difference in the power of the test;[8] (5) applying the revealed price preference method of Deb *et al.* (2022) to test consistency with augmented utility maximization;[9] (6) measuring downward-sloping demand to proxy individual rationality level (Echenique *et al.*, forthcoming).[10] Second, for risk preference measured in the lab, we apply recent techniques in revealed preference to further test whether the utility function respects to first order stochastic dominance (Nishimura *et al.*, 2017), expected utility (Polisson *et al.*, 2020), and downward-sloping demand (Echenique *et al.*, forthcoming). Table 3 presents the pairwise correlation test results. Again, all correlation coefficients are close to zero and statistically insignificant. We find similar results when using HMI, MPI and MCI in Table A4. We summarize these findings below.

Result 1. *The correlation of rationality between food and risk preferences is low.*

[7]Specifically, we obtain price P_{It}, by averaging prices weighted by quantity: $P_{It} = \frac{\sum_{i \in I} e_{it}}{Q_{It}}$, where e_{it} is all consumers' expenditure for item i in category I.

[8]Consumers may be presented with different budget sets and have different power for testing GARP violations. We follow Beatty and Crawford (2011) and compute Selten's score by subtracting Bronars' index from the CCEI

[9]Since only choices for vegetable and meat are observed, we do not know the full sets for consumption. This takes into account the impact of the prices of observed goods on consumption of the unobserved goods.

[10]Roughly speaking, prices and quantities should be inversely related. We calculate the correlation coefficient between $log(x_2/x_1)$ and $log(p_2/p_1)$ for each subject in the datasets

Table 3: Correlations of CCEI in Study 1: Food vs. Risk

	Risk			
	Original	*FOSD*	*EU*	*Downward*
Food				
Original	0.001	-0.025	-0.030	-0.013
	(0.978)	(0.419)	(0.334)	(0.676)
Same price	-0.001	0.026	0.025	0.026
	(0.976)	(0.403)	(0.406)	(0.398)
Three year	0.007	-0.003	-0.005	0.013
	(0.809)	(0.920)	(0.863)	(0.669)
Three goods	-0.009	0.009	0.007	0.018
	(0.776)	(0.756)	(0.820)	(0.561)
Selten scores	0.002	-0.015	-0.021	0.019
	(0.955)	(0.615)	(0.483)	(0.524)
GAPP	-0.015	0.003	0.001	0.030
	(0.627)	(0.926)	(0.982)	(0.332)
Downward	0.013	-0.006	-0.000	-0.022
	(0.677)	(0.837)	(0.994)	(0.465)

Note: Spearman's ρ are reported. P-value in parentheses.

3.3.3 Understanding inconsistency between food and risk preferences

The observed low consistency in rationality suggests that different decision-making mechanisms may underlie these two preferences. Therefore, we conduct regression analyses to examine which factors are related to the rationality level for each type of preferences.

We examine how individuals' demographics are correlated with rationality measures (Choi *et al.*, 2014; Echenique *et al.*, 2011). Table 4 presents OLS regression analyses, in which the dependent variables are the CCEI score for food preference (Column 1) and risk preference (Column 2). Independent variables are female, income, education, age dummies and family size.

Table 4: OLS Regressions for Demographics and CCEI in Study 1

	Food	Risk
	(1)	(2)
Female	0.017***	-0.006
	(0.005)	(0.006)
Medium Income	0.015***	0.022***
	(0.005)	(0.007)
High income	0.023***	0.046***
	(0.006)	(0.007)
Medium education	0.001	0.001
	(0.005)	(0.006)
High education	0.000	0.008
	(0.006)	(0.007)
Medium age	-0.008	-0.014
	(0.010)	(0.013)
Elder age	-0.028***	-0.052***
	(0.011)	(0.014)
Family size	0.011***	-0.003
	(0.002)	(0.003)
N	1073	1073
R^2	0.080	0.088

Note: Medium (High) income indicates whether family income is 5000-10000 (more than 10000) RMB per month. Medium (High) education is high school (above high school). Medium (Elder) age refers to those born 1970-1989 (before 1969). Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

First, the CCEI score for both preferences is significantly positively correlated with family income and their correlation with age is negative and significant, which is in line with previous studies (Choi *et al.*, 2014; Echenique *et al.*, 2011). We also find that the size of the coefficients for high-income (elder-age) dummies are significantly larger (smaller) for risk preference than for food preference, which suggests that richer and younger people are better at making high-quality decisions for risk tasks. Moreover, females and subjects with larger family size have higher scores for food preference, but not risk preference. Altogether, these results represent suggestive evidence that individual demographics correlate with individuals' decision-making quality in these two types of tasks.[11]

Moreover, we examine whether and to what extent individuals' shopping experience and

[11]We obtain similar results for HMI, MPI and MCI (Table A5).

cognitive ability affects their rationality level and report the results in Table 5. Again, the dependent variable is the CCEI score for food preference (Columns 1-4) and risk preference (Columns 5-8). For independent variables, we include monthly expenditure on vegetables and meat from 2019 to 2020, shopping frequency using monthly days of shopping for vegetables and meat from 2019 to 2020, Raven's IQ test scores and Big Five personality traits. We also control for aforementioned demographic variables.

First, subjects with more expenditure and higher shopping frequency have significantly higher CCEI scores for food preference, but not for risk preference. In contrast, subjects with higher IQ scores have higher CCEI scores for risk preference, but not for food preference. This suggests that rationality for food preference in the supermarket can be improved by shopping experiences, while rationality for risk preference in the lab may require more abstract reasoning. We also observe a significantly positive relationship between conscientiousness and CCEI score for food preference, which is consistent with prior findings,[12] and a significantly negative relationship between neuroticism and CCEI score for risk preference.[13] Taken together, the observed distinct correlates are in line with the low consistency between rationality measures for food and risk preferences. We also find similar results for HMI, MPI and MCI (Table A6). We summarize these findings below.

Result 2. *Rationality for both food and risk preferences are positively (negatively) correlated with family income (age). The rationality for food preference is correlated with shopping experience and conscientiousness, while the rationality for risk preference is correlated with cognitive ability and neuroticism.*

[12]Conscientiousness is associated with dependability, organizational sills, perseverance, and achievement oriented thinking, which contributes to prudence with respect to food purchases and more stable preferences (Borghans *et al.*, 2008)

[13]Neuroticism is associated with anxiety, worry, anger, and insecurity. People with high neuroticism may become more nervous when they face a risk decision task, which results in a decline in the quality of decision-making (Byrne *et al.*, 2015)

Table 5: OLS Regressions for Behavioral Factors and CCEI in Study 1

	Food				Risk			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expenditure	0.020***			0.013***	0.003			0.003
	(0.004)			(0.005)	(0.006)			(0.006)
Frequency		0.221***		0.168***		0.049		0.038
		(0.048)		(0.053)		(0.062)		(0.066)
IQ			-0.001	-0.001			0.012***	0.012***
			(0.001)	(0.001)			(0.002)	(0.002)
Conscientiousness				0.008***				-0.004
				(0.002)				(0.003)
Extraversion				-0.001				-0.003
				(0.002)				(0.003)
Agreeableness				0.001				-0.003
				(0.002)				(0.003)
Openness				0.003				-0.004
				(0.002)				(0.003)
Neuroticism				-0.000				-0.013***
				(0.002)				(0.003)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1073	1073	1073	1073	1073	1073	1073	1073
R^2	0.097	0.098	0.081	0.117	0.088	0.088	0.143	0.167

Note: Control variables include gender, income, education, age dummies and family size. Standard errors are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4 Study 2

4.1 Lab: Risk/Social/Food Preferences

We show in Study 1 that the correlation between rationality scores for risky decisions in the lab and food decisions in the field is low. The observed low correlation could be due to various reasons such as types of decisions, namely, risky decisions versus food decisions, or lab versus field settings. In particular, in the field setting, scanner data may not accurately reveal rationality of food preference, because the budget constraints differ across individuals leading to different power of the test (Beatty and Crawford, 2011; Dean and Martin, 2016), and the underlying preferences may also change over a time period of the two years. We conduct a second experiment (Study 2) to partially address this concern. In addition to risky decisions in the lab and food decisions in the field as in Study 1, we also measure the rationality of social and food decisions in the lab.

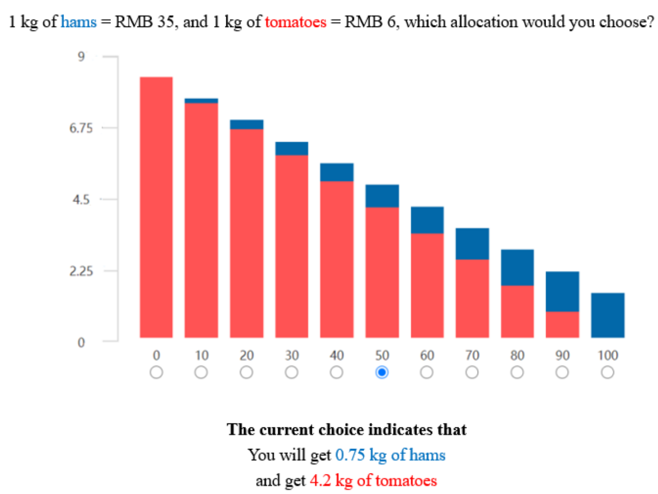


Figure 5: Screenshot for Food Preference Task

To measure rationality in food preference in the lab, we construct a purchasing environment in which subjects are endowed with a fixed amount of expenditure (RMB 50). They allocate the RMB 50 between a specific type of vegetable (tomato) and a specific type of meat (ham). They make allocation decisions for 22 rounds. The price level for each product is chosen from the range of actual prices in 2020-2021. Figure 5 presents a sample screenshot for the food task.

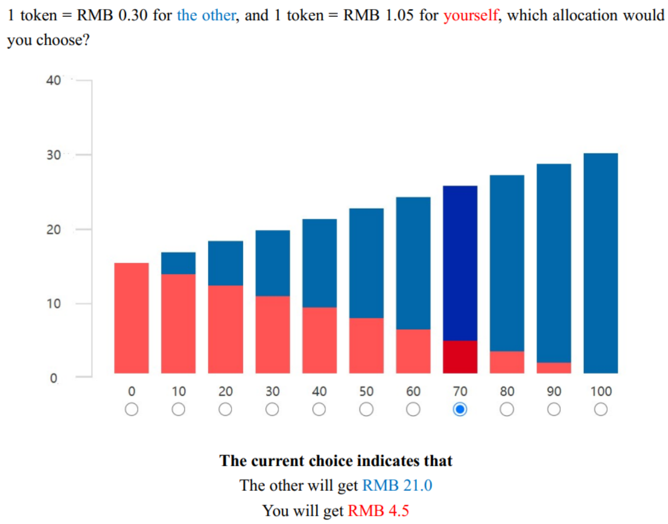


Figure 6: Screenshot for Social Preference Task

To measure rationality in social preference, we use a modified dictator game (Andreoni and Miller, 2002; Fisman *et al.*, 2007). Subjects allocate 100 tokens between themselves and another randomly matched subject across 22 rounds. In every round, each token has a different cash value for the subject and their matched subject. Figure 6 presents a sample screenshot. We also include the same risk task as in Study 1.

We conducted the experiment from December 2021 to February 2022. Appendix C details instructions for the experiment. In total, 305 customers participated in the survey. We presented the risk, social and food tasks in random order to subjects and end the experiment with the same questionnaire as in Study 1. Subjects spent on average 30.1 minutes to complete the experiments and received an average payment of RMB 55.1.

We also quantify the rationality score for food preferences using supermarket transaction data. Since the consumption frequency for participants in Study 2 is not as high as in Study 1, our sample selection procedure is less strict, i.e., we restrict our analyses to 174 consumers who bought both meat and vegetables for at least 7 months between 2019 and 2020. As a consequence, our sample size is 305 for the results without rationality in the field, and is 174 for the results involving rationality in the field. Moreover, to address the problem whereby some consumers does not have a purchase record in a certain month, we assume that all households are presented with the same price for each goods category within a month (Dean and Martin, 2016; Echenique *et al.*, 2011). Specifically, we obtain the price for each consumption category in month t, denoted by P_{It}, by averaging the prices weighted by quantity:$P_{It} = \frac{\sum_{i \in I} e_{it}}{\sum_{i \in I} q_{it}}$ where e_{it} and q_{it} is all consumers' expenditure and quantity of consumption for item i in category I.

4.2 Study 2: Results

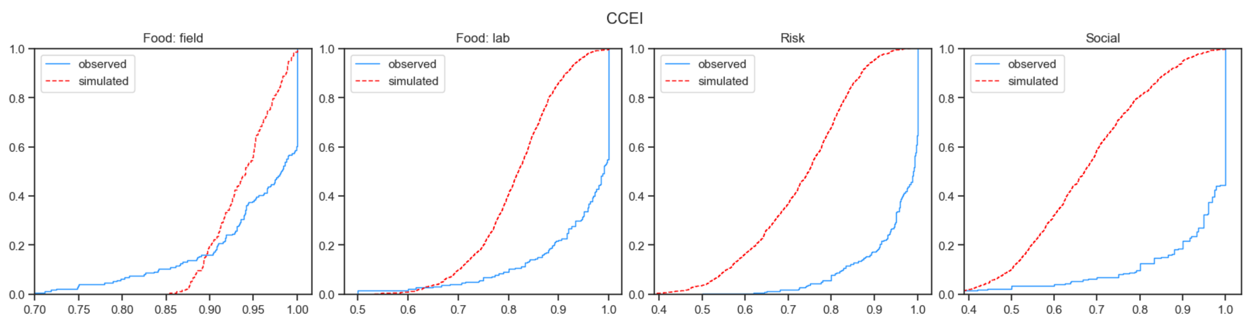


Figure 7: Rationality Index across Settings

Figure 7 presents the cumulative distribution of CCEI scores (solid line) for food preference with supermarket consumption, food preference in the lab, and risk and social preferences (see also summary statistics in Tables B2-B5). We observe a high level of rationality in risk, social and food preferences, which is consistent with results in Study 1. We also plot Bronars' index (dashed line) to capture the rationality level for a subject who make choices randomly (Beatty and Crawford, 2011; Bronars, 1987; Dean and Martin, 2016), which suggests that the chosen set of parameters has sufficient power to detect GARP violations.

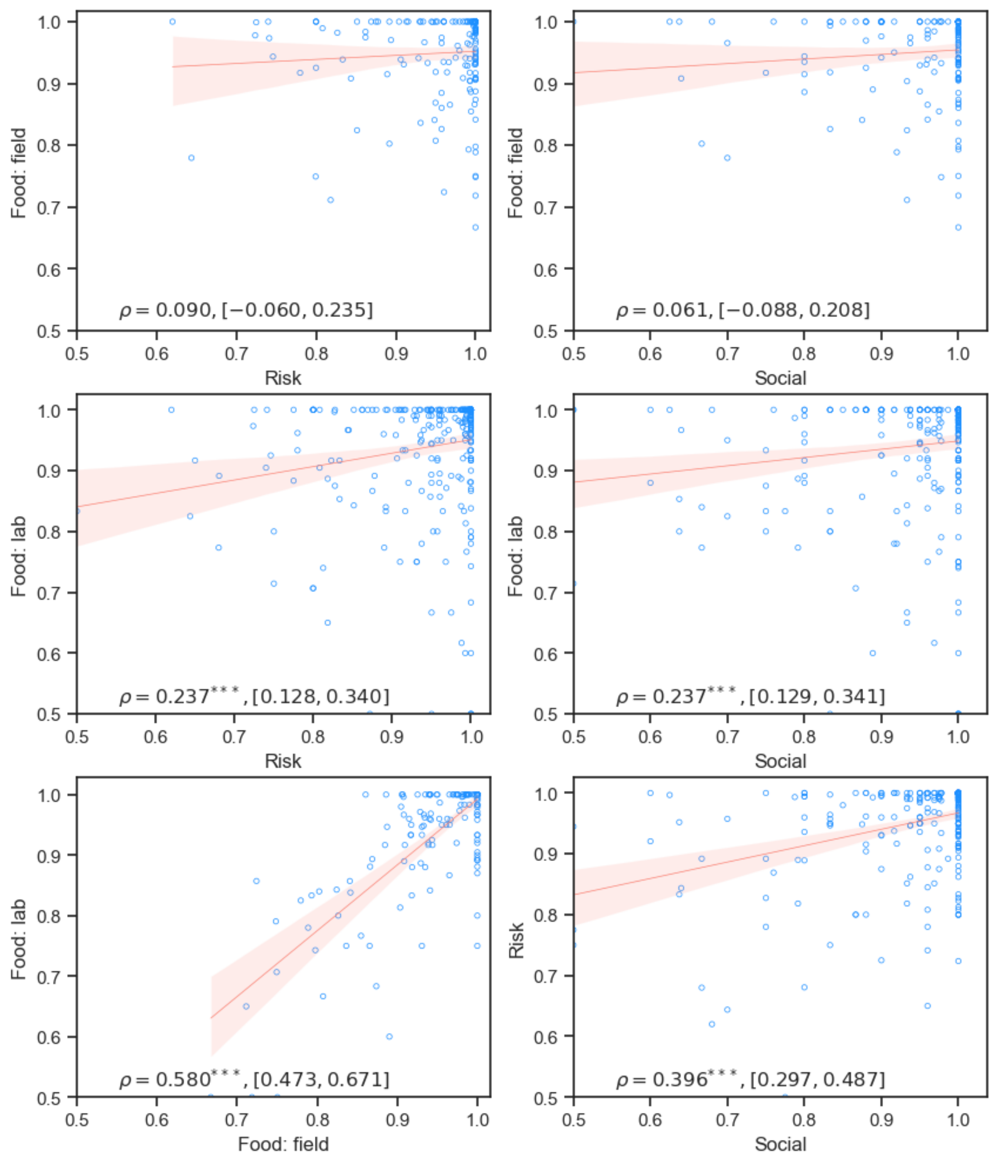


Figure 8: Correlations of CCEI in Study 2: Food vs. Risk vs. Social Preferences

Next, we test pairwise correlations for rationality scores across domains (Table 1 and Figure 8). First, consistent with Study 1, we find that there is no significant correlation of CCEI scores between risk preference in the lab and food preference in the field. Similarly, we also find low correlation between social preference in the lab and food preference in the field. However, the rationality between risk and social preferences in the lab is significantly correlated (Spearman correlation coefficient: 0.396, $p < 0.01$), suggesting that consumers with higher rationality score in risk preference tend to exhibit a higher rationality level in social preference in the lab. Moreover, individuals exhibit internal consistency for food preference in the lab and in the field (Spearman correlation coefficient: 0.580, $p < 0.01$). The rationality score of food preference in the lab is weakly correlated with both that of risk preference (Spearman correlation coefficient: 0.237, $p < 0.01$) and social preference (Spearman correlation coefficient: 0.237, $p < 0.01$). Additionally, by applying the alternative measures discussed in Study 1, we find that the results here are robust (Tables B6-B8). This leads to our next result.

Result 3. *(a) The correlation of rationality is high between risk and social preference in the lab, as well as between food preference in the lab and in the field. (b) The correlation of rationality is low between risk/social preference in the lab and food preference in the field, and it is moderate between risk/social preference in the lab and food preference in the lab.*

We conduct the same OLS regressions as in Study 1 to examine the correlation between rationality scores and individual characteristics. The overall observations are similar to those in Study 1 (Table 6). First, total expenditure matters for rationality scores for food preference, but not for social and risk preferences. Second, IQ has a significantly positive association with both risk and social preferences, but not with food preference. Third, there is a significantly positive relationship between conscientiousness and rationality scores for food preference, and a negative relationship between neuroticism and CCEI scores for risk preference.[14] Lastly, analyses of demographic variables are similar to Study 1 (Tables B9 and B10), which demonstrates that family income is positively correlated with rationality scores measured in different settings.

[14]Similar results for regressions of HMI, MPI and MCI are reported in Tables B11 and B12.

Table 6: OLS Regressions for Behavioral Factors and CCEI in Study 2

	Food: field	Food: lab	Risk	Social
	(1)	(2)	(3)	(4)
Expenditure	0.039**	0.051*	-0.021	-0.042
	(0.018)	(0.026)	(0.019)	(0.036)
Frequency	-0.147	0.287	-0.108	0.264
	(0.169)	(0.242)	(0.174)	(0.339)
IQ	-0.008	-0.035	0.043**	0.118***
	(0.020)	(0.028)	(0.021)	(0.040)
Conscientiousness	0.015***	0.022***	-0.001	0.002
	(0.005)	(0.007)	(0.005)	(0.010)
Extraversion	0.002	-0.005	-0.000	-0.013
	(0.005)	(0.008)	(0.006)	(0.011)
Agreeableness	0.000	-0.008	0.005	-0.000
	(0.006)	(0.008)	(0.006)	(0.011)
Openness	0.000	0.010	-0.001	-0.003
	(0.005)	(0.008)	(0.006)	(0.011)
Neuroticism	0.002	-0.001	-0.012**	-0.003
	(0.005)	(0.008)	(0.006)	(0.011)
Control	Yes	Yes	Yes	Yes
N	174	174	174	174
R^2	0.218	0.206	0.259	0.186

Note: Control variables are gender, income, education, age dummies and family size. In order measure expenditure and rationality in the field, we restrict our analyses to 174 consumers who bought both meat and vegetables for at least 7 months in the two-year period between 2019 and 2020. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Taking the results in Studies 1 and 2 together, we observe that the rationality in food preferences is correlated with the same sets of demographic and behavioral variables, namely purchasing experience and conscientiousness, regardless of whether this is based on the scanner data set or measured in the lab. The correlates for rationality in risk and social preferences in the lab are cognitive ability and neuroticism. These factors play different roles across settings of choices and in part lead to the observed differences in consistency across rationality measures.

5 Discussion

We show that the consistency of rationality measures depends on either the lab or field setting as well as on the type of decisions. Our observations suggest that while rationality scores can be inferred in various environments, they may not always or solely reflect the general ability to make choices according to one's preferences. This message is shared by studies inferring preferences from observable choices. It is often challenging to measure and separate the influences of unobservable factors, such as budget constraints and heuristics. For example, Dean and Sautmann (2021) shows that choices in the experiments may not fully reflect true time preference, since it can also be influenced by financial shocks and budget constraints. Enke and Graeber (2019) finds that the commonly observed S-shaped probability-weighting function in prospect theory can in part be due to cognitive uncertainty, whereby subjects exhibit a tendency to choose the cognitive default, e.g., switching around the middle of the choice list. In a similar vein, rationality scores inferred from observable choices may be "as if" to some degree. Below, we discuss several potential issues with the identification and measurement of rationality which may help understand the observed consistency in our study.

Formation of preferences. The formation of preferences can differ across environments. Plott (1996) proposes the notion of preference discovery, whereby rationality can be understood as a discovery process in which subjects learn their own needs through a process of reflection and practice. Deming (2021) suggests the distinction between the decision-making ability of doing problem-solving task and routine task. In this regard, decision maker as a subject in the lab is more likely to perform a problem-solving task whereas decision maker as a consumer is more likely to do a routine task. In the lab, when subjects are presented with risky decisions and social decisions, they need to learn the abstract experimental rules, since these rules are not part of their daily experience. In this regard, rationality in the lab may be related to the ability to "discover" one's preference in the new choice environment and "solve" the choice problem accordingly. Consistent with this view, we find that the rationality scores are lower for the first half than the second half of the risky and social decisions.[15] Because the rationality scores of risky and social decisions in the lab may in part capture this ability, we observe that they are highly correlated and both are positively correlated with cognitive ability. Choi *et al.* (2014) shows that rationality scores in the lab are positively correlated with household wealth after controlling for unobserved constraints, preferences, beliefs, and cognitive ability.

By contrast, when consumers make food decisions in daily life, they form their preferences through their purchase and consumption experiences (Crawford, 2010). Food decisions may thus appear to be more routine for the consumers with more experience. For example, List (2003) provides field evidence showing that individuals with more trading experience exhibit less endowment effect and behave closer to neoclassical predictions, and suggests that consumers can learn to overcome the endowment effect. Related to this point, we show that the consumers' experiences also matter for rationality; those with higher purchasing frequency tend to have higher rationality scores in the field. Moreover, their daily experiences may further carry over when they make food decisions in the lab, which may help explain the high correlation with food decisions in the field.[16]

Moreover, the food preferences are often influenced by various contextual factors, including meal vs. non-meal time, working vs. non-working days, and discounted vs. non-discounted purchases. We examine differences in preferences between meal time and non-meal time purchases in Study 1. We refer purchases from 10:00 to 14:00 and from 16:00 to 19:00 as

[15]The average CCEI score is 0.969, 0.974 and 0.960 for the first half of the choices and 0.981, 0.982 and 0.980 for the second half of the choices for risky decisions in Study 1, risky decisions in Study 2, and social decisions in Study 2, respectively. We also compute the CCEI in Choi *et al.* (2007) and Choi *et al.* (2014) and find that the average CCEI score is 0.960 and 0.937 for the first half and 0.986 and 0.944 for the second half of the experimental data, respectively. By contrast, we do not observe such difference in food decisions. Namely, the average CCEI score is 0.967 for the first half and 0.969 for the second half for food decisions in Study 2. This suggests that consumers do not need to "discover" their food preferences as they are familiar with these decisions.

[16]Even if subjects do have clear preferences in mind, rationality can also be influenced by preferences. For example, in the social decision tasks, a purely selfish decision maker can easily maximize her utility by giving nothing to the recipient, while a decision maker with selfish, equity and efficiency considerations may find it hard to trade off various motives, which leads to a lower rationality score.

meal time purchases and the rest as non-meal time purchases. We compute the rationality scores separately using the datasets from meal time and non-meal time purchases, as well as the combined dataset which randomly mixes 50 percent of the observations from each two datasets (Castillo and Freer, 2018; Miao *et al.*, 2021). We find that the average CCEI score is 0.92 and 0.932 for meal time purchases and non-meal time purchases, respectively, but it is 0.878 for the combined data, which is significantly lower than that of the two separate datasets ($p<0.001$, two-sided paired t-test). We observe a similar pattern for working vs. non-working days, and discounted vs. non-discounted purchases. Lower CCEI scores for the combined dataset suggest that food preferences are subjective to these contextual influences.

Budget constraints. Budget constraints can also differ across environments.[17] In the lab, subjects are presented with well-defined budget lines using a simple interface. By contrast, consumption data often lack the power to detect GARP violations, because budget lines may not sufficiently cross (Blundell *et al.*, 2003, 2008; Chambers and Echenique, 2016, Chapter 5). While our scanner data have the power to detect GARP violations, and we examine several alternative measures, it is generally difficult to construct budget lines; in particular, perceived prices can be affected by the consumer's attention and memory (Bordalo *et al.*, 2020). For example, a consumer may not notice the price change when the new price is not too different from the price in her memory database, and may perceive some products as very cheap when the new price is sufficiently lower than the price in her memory database.

Heuristic rules. In addition to preference and budget constraints, different environments may induce decision makers to use different heuristic rules, especially when facing difficult decisions (Gigerenzer and Todd, 1999; Simon, 1979; Tversky and Kahneman, 1974). In the environment of budgetary choices, Choi *et al.* (2006) shows that subjects' portfolio choices can be explained by some simple rules, such as a diversification heuristic—by allocating

[17]Varian (1988) show that the utility maximization hypothesis is not falsifiable if we can not observe consumer's choice for all goods, and van Bruggen and Heufer (2017) show that it is possible if unobserved prices and expenditures remain constant. In this regard, it is more likely for the decision makers to see the observed prices and expenditures as constant in the lab than in the field.

the same number of points to the two accounts. Halevy and Mayraz (2022) designs simple investment rules for selecting portfolios and show that most of the subjects prefer to make allocations through the rule-based interface. It is possible that in the abstract environment of risky and social decisions in the lab, subjects find the decisions to be difficult and adopt a set of heuristic rules to simplify their choices. In food decisions, decision makers may use different heuristic rules or habits formed in their daily life (Havranek *et al.*, 2017). Thaler and Shefrin (1981) proposes the importance of heuristics in consumer choice. For example, consumers are more motivated to save \$5 on a \$15 item than to save \$5 on a \$125 item. These heuristics may drive rationality scores in consumer choice.

Summary. Whereas the revealed preference approach provides a unified and powerful framework for analyzing choice behavior, preferences and rationality revealed from observable choices need not be consistent across environments. Choi *et al.* (2014) discusses decision-making quality observed in the field and in the lab, and suggest that measuring decision-making quality in the field is difficult because decision makers "might have different preferences over the same outcomes, or face different but unobserved incentives and constraints, or have different information, or hold different beliefs." They further propose the departure from GARP in the lab as rationality scores to measure decision-making quality. By combining experimental data on risky, social, and food choices with customers' scanner data, our study is the first to examine the consistency of rationality measures. In accordance with the view in Choi *et al.* (2014), we show that the consistency between risky/social decisions in the lab and food decisions in the field is indeed low. Moreover, we observe a high consistency of rationality measures between risky and social decisions in the lab, which suggests that both of the measures capture some general decision-making quality, for example, the ability to learn and express one's preference in an abstract environment. We also observe a high consistency between food decisions in the lab and in the field, which may reflect the general influences of preferences and habits formed from daily experiences of the consumers. We leave for the future studies to advance identification and measurement challenge toward a better understanding of rationality across environments.

References

Afriat SN (1967). “The construction of utility functions from expenditure data.” *International Economic Review*, **8**(1), 67–77.

Afriat SN (1972). “Efficiency estimation of production functions.” *International Economic Review*, **13**(3), 568–598.

Ahn D, Choi S, Gale D, Kariv S (2014). “Estimating ambiguity aversion in a portfolio choice experiment.” *Quantitative Economics*, **5**(2), 195–223.

Andreoni J, Miller J (2002). “Giving according to GARP: An experimental test of the consistency of preferences for altruism.” *Econometrica*, **70**(2), 737–753.

Augenblick N, Niederle M, Sprenger C (2015). “Working over time: Dynamic inconsistency in real effort tasks.” *The Quarterly Journal of Economics*, **130**(3), 1067–1115.

Banks J, Carvalho L, Perez-Arce F (2019). “Education, decision making, and economic rationality.” *Review of Economics and Statistics*, **101**(3), 428–441.

Barseghyan L, Molinari F, O’Donoghue T, Teitelbaum JC (2013). “The nature of risk preferences: Evidence from insurance choices.” *American economic review*, **103**(6), 2499–2529.

Barseghyan L, Prince J, Teitelbaum JC (2011). “Are risk preferences stable across contexts? Evidence from insurance data.” *American Economic Review*, **101**(2), 591–631.

Beatty TK, Crawford I (2011). “How demanding is the revealed preference approach to demand?” *American Economic Review*, **101**(6), 2782–95.

Blundell R, Browning M, Crawford I (2003). “Nonparametric Engel curves and revealed preference.” *Econometrica*, **71**(1), 205–240.

Blundell R, Browning M, Crawford I (2008). “Best nonparametric bounds on demand responses.” *Econometrica*, **76**(6), 1227–1262.

Bordalo P, Gennaioli N, Shleifer A (2020). “Memory, attention, and choice.” *The Quarterly Journal of Economics*, **135**(3).

Borghans L, Duckworth AL, Heckman JJ, Ter Weel B (2008). “The economics and psychology of personality traits.” *Journal of Human Resources*, **43**(4), 972–1059.

Bronars SG (1987). “The power of nonparametric tests of preference maximization.” *Econometrica*, **55**(3), 693–698.

Burks S, Carpenter J, Götte L, Rustichini A (2012). "Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment." *Journal of Economic Behavior & Organization*, **84**(1), 308–320.

Byrne KA, Silasi-Mansat CD, Worthy DA (2015). "Who chokes under pressure? The Big Five personality traits and decision-making under pressure." *Personality and Individual Differences*, **74**, 22–28.

Camerer C (2011). "The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List." *Working Paper*.

Cappelen AW, Kariv S, Sørensen EØ, Tungodden B (2021). "The development gap in economic rationality of future elites." *Working Paper*.

Carvalho L, Meier S, Wang SW (2016). "Poverty and economic decision-making: Evidence from changes in financial resources at payday." *American Economic Review*, **106**(2), 260–84.

Carvalho L, Olafsson A, Silverman D (2019). "Misfortune and mistake: the Financial conditions and decision-making Ability of high-cost loan borrowers." *Working Paper*.

Castillo M, Freer M (2018). "Revealed differences." *Journal of Economic Behavior & Organization*, **145**, 202–217.

Chambers CP, Echenique F (2016). *Revealed preference theory*. Cambridge University Press.

Chapman J, Dean M, Ortoleva P, Snowberg E, Camerer C (2023). "Econographics." *Journal of Political Economy Microeconomics*, **1**(1), 000–000.

Choi S, Fisman R, Gale D, Kariv S (2006). "Substantive and procedural rationality in decisions under uncertainty." *Working Paper*.

Choi S, Fisman R, Gale D, Kariv S (2007). "Consistency and heterogeneity of individual behavior under uncertainty." *American Economic Review*, **97**(5), 1921–1938.

Choi S, Kariv S, Müller W, Silverman D (2014). "Who is (more) rational?" *American Economic Review*, **104**(6), 1518–50.

Cohn A, Maréchal MA (2018). "Laboratory measure of cheating predicts school misconduct." *The Economic Journal*, **128**(615), 2743–2754.

Crawford I (2010). "Habits revealed." *The Review of Economic Studies*, **77**(4), 1382–1402.

Crawford I, De Rock B (2014). "Empirical revealed preference." *Annu. Rev. Econ.*, **6**(1), 503–524.

Currie J, MacLeod WB (2017). "Diagnosing expertise: Human capital, decision making, and performance among physicians." *Journal of Labor Economics*, **35**(1), 1–43.

Dai Z, Galeotti F, Villeval MC (2018). "Cheating in the lab predicts fraud in the field: An experiment in public transportation." *Management Science*, **64**(3), 1081–1100.

Dean M, Martin D (2016). "Measuring rationality with the minimum cost of revealed preference violations." *Review of Economics and Statistics*, **98**(3), 524–534.

Dean M, Ortoleva P (2019). "The empirical relationship between nonstandard economic behaviors." *Proceedings of the National Academy of Sciences*, **116**(33), 16262–16267.

Dean M, Sautmann A (2021). "Credit constraints and the measurement of time preferences." *Review of Economics and Statistics*, **103**(1), 119–135.

Deb R, Kitamura Y, Quah JKH, Stoye J (2022). "Revealed Price Preference: Theory and Empirical Analysis." *The Review of Economic Studies*.

Deming DJ (2021). "The growing importance of decision-making on the job." *Working Paper*.

Dohmen T, Falk A, Huffman D, Sunde U, Schupp J, Wagner GG (2011). "Individual risk attitudes: Measurement, determinants, and behavioral consequences." *Journal of the European Economic Association*, **9**(3), 522–550.

Echenique F (2020). "New developments in revealed preference theory: Decisions under risk, uncertainty, and intertemporal choice." *Annual Review of Economics*, **12**, 299–316.

Echenique F (2021). "On the meaning of the critical cost efficiency index." *Working Paper*.

Echenique F, Lee S, Shum M (2011). "The money pump as a measure of revealed preference violations." *Journal of Political Economy*, **119**(6), 1201–1223.

Echenique F, Saito K, Imai T (forthcoming). "Approximate expected utility rationalization." *Journal of the European Economic Association*.

Einav L, Finkelstein A, Pascu I, Cullen MR (2012). "How general are risk preferences? Choices under uncertainty in different domains." *American Economic Review*, **102**(6), 2606–38.

Enke B, Graeber T (2019). "Cognitive uncertainty." *Working Paper*.

Fisman R, Kariv S, Markovits D (2007). "Individual preferences for giving." *American Economic Review*, **97**(5), 1858–1876.

Galizzi MM, Navarro-Martinez D (2019). "On the external validity of social preference games: a systematic lab-field study." *Management Science*, **65**(3), 976–1002.

Georganas S, Healy PJ, Weber RA (2015). "On the persistence of strategic sophistication." *Journal of Economic Theory*, **159**, 369–400.

Gigerenzer G, Todd PM (1999). “Fast and frugal heuristics: The adaptive toolbox.” In “Simple heuristics that make us smart,” pp. 3–34. Oxford University Press.

Halevy Y, Mayraz G (2022). “Identifying rule-based rationality.” *The Review of Economics and Statistics*, pp. 1–44.

Halevy Y, Persitz D, Zrill L (2018). “Parametric recoverability of preferences.” *Journal of Political Economy*, **126**(4), 1558–1593.

Havranek T, Rusnak M, Sokolova A (2017). “Habit formation in consumption: A meta-analysis.” *European Economic Review*, **95**, 142–167.

Houtman M, Maks J (1985). “Determining all maximal data subsets consistent with revealed preference.” *Kwantitatieve Methoden*, **19**(1), 89–104.

Kahneman D (2003). “Maps of bounded rationality: Psychology for behavioral economics.” *American Economic Review*, **93**(5), 1449–1475.

Kim HB, Choi S, Kim B, Pop-Eleches C (2018). “The role of education interventions in improving economic rationality.” *Science*, **362**(6410), 83–86.

Levitt SD, List JA (2007a). “Viewpoint: On the generalizability of lab behaviour to the field.” *Canadian Journal of Economics-Revue Canadienne D Economique*, **40**, 347370.

Levitt SD, List JA (2007b). “What do laboratory experiments measuring social preferences reveal about the real world?” *Journal of Economic Perspectives*, **21**(2), 153–174.

Levitt SD, List JA, Sadoff SE (2011). “Checkmate: Exploring backward induction among chess players.” *American Economic Review*, **101**(2), 975–90.

Li W, Leng Z, Yi J, Zhong S (2023). “A multifaceted poverty-reduction program has economic and behavioral consequences.” *Proceedings of the National Academy of Sciences*, **forthcoming**.

List JA (2003). “Does market experience eliminate market anomalies?” *The Quarterly Journal of Economics*, **118**(1), 41–71.

Miao B, Yang S, Zhong S (2021). “Revealed heterogeneity in risk preference.” *Working Paper*.

Nishimura H, Ok EA, Quah JKH (2017). “A comprehensive approach to revealed preference theory.” *American Economic Review*, **107**(4), 1239–63.

Plott CR (1996). “Rational individual behavior in markets and social choice processes: The discovered preference hypothesis.” In “The Rational Foundations of Economic Behavior,” Macmillan.

Polisson M, Quah JKH (2013). "Revealed preference in a discrete consumption space." *American Economic Journal: Microeconomics*, **5**(1), 28–34.

Polisson M, Quah JKH (2022). "Rationalizability, cost-rationalizability, and Afriat's efficiency index." *Working Paper*.

Polisson M, Quah JKH, Renou L (2020). "Revealed preferences over risk and uncertainty." *American Economic Review*, **110**(6), 1782–1820.

Rubinstein A (2016). "A typology of players: Between instinctive and contemplative." *The Quarterly Journal of Economics*, **131**(2), 859–890.

Samuelson PA (1938). "A note on the pure theory of consumer's behaviour." *Economica*, **5**(17), 61–71.

Selten R (1991). "Properties of a measure of predictive success." *Mathematical Social Sciences*, **21**(2), 153–167.

Simon HA (1979). "Rational decision making in business organizations." *The American Economic Review*, **69**(4), 493–513.

Stango V, Zinman J (2022). "We are all behavioral, more or less: Measuring and using consumer-level behavioral sufficient statistics." *The Review of Economic Studies*.

Thaler RH (2016). "Behavioral economics: Past, present, and future." *American Economic Review*, **106**(7), 1577–1600.

Thaler RH, Shefrin HM (1981). "An economic theory of self-control." *Journal of Political Economy*, **89**(2), 392–406.

Tversky A, Kahneman D (1974). "Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty." *Science*, **185**(4157), 1124–1131.

van Bruggen P, Heufer J (2017). "Afriat in the lab." *Journal of Economic Theory*, **169**, 546–550.

Varian HR (1982). "The Nonparametric Approach to Demand Analysis." *Econometrica*, **50**(4), 945–973.

Varian HR (1988). "Revealed preference with a subset of goods." *Journal of Economic Theory*, **46**(1), 179–185.

Varian HR (1990). "Goodness-of-fit in optimizing models." *Journal of Econometrics*, **46**(1-2), 125–140.

Weber EU, Blais AR, Betz NE (2002). "A domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors." *Journal of Behavioral Decision Making*, **15**(4), 263–290.

Online Appendices

A Study 1

Table A1: Demographics of Subjects in Study 1

Demographic Variable	Category	Percent
Gender	*Female*	75.1%
	Male	24.9%
Age	*Before 1969*	24.6%
	In 1970-1989	70.8%
	After 1990	4.6%
Family Monthly Income	*<5000*	23.6%
	5000-10000	40.8%
	>10000	35.6%
Education	*High school below*	46.3%
	High school	29.9%
	High school above	23.8%
Family Size	*One*	1.6%
	Two	9.3%
	Three	38.5%
	Four	30.5%
	>Four	20.1%

Table A2: Rationality Index in Food preference

	Mean	SD	Min	Median	Max
Observed					
CCEI	0.941	0.070	0.597	0.962	1
HMI	0.924	0.053	0.708	0.917	1
MPI	0.059	0.054	0	0.047	0.300
MCI	0.004	0.007	0	0.002	0.061
Obs: 1,073					
Bronars					
CCEI	0.848	0.034	0.74	0.848	0.947
HMI	0.859	0.026	0.769	0.860	0.932
MPI	0.125	0.022	0.061	0.124	0.220
MCI	0.018	0.008	0.002	0.016	0.060
Obs: 1,073					

Table A3: Rationality Index in Risk preference

	Mean	SD	Min	Median	Max
Observed					
CCEI	0.943	0.089	0.350	0.987	1
HMI	0.883	0.092	0.636	0.909	1
MPI	0.041	0.054	0	0.020	0.379
MCI	0.009	0.023	0	0.001	0.321
Obs: 1,073					
Bronars					
CCEI	0.709	0.122	0.352	0.723	0.967
HMI	0.742	0.073	0.500	0.727	0.955
MPI	0.159	0.045	0.042	0.158	0.320
MCI	0.088	0.058	0.003	0.076	0.354
Obs: 1,000					

Table A4: Correlations of Rationality between field and risk preference

	Risk									
	CCEI		HMI		MPI		MCI		*EU*	*Downward*
	Original	*FOSD*	*Original*	*FOSD*	*Original*	*FOSD*	*Original*	*FOSD*		
Food										
Original	0.001	-0.025	0.051	-0.009	0.015	-0.020	-0.002	-0.033	-0.030	-0.013
	(0.978)	(0.419)	(0.094)	(0.771)	(0.623)	(0.510)	(0.943)	(0.277)	(0.334)	(0.676)
Same price	-0.001	0.026	0.006	0.043	0.009	0.035	-0.002	0.008	0.025	0.026
	(0.976)	(0.403)	(0.834)	(0.156)	(0.779)	(0.251)	(0.939)	(0.798)	(0.406)	(0.398)
Three year	0.007	-0.003	-0.024	-0.002	0.026	0.009	-0.012	-0.021	-0.005	0.013
	(0.809)	(0.920)	(0.426)	(0.937)	(0.397)	(0.778)	(0.702)	(0.502)	(0.863)	(0.669)
Three goods	-0.009	0.009	-0.008	0.023	-0.004	0.017	0.007	0.018	0.007	0.018
	(0.776)	(0.756)	(0.781)	(0.459)	(0.904)	(0.578)	(0.824)	(0.562)	(0.820)	(0.561)
Selten scores	-0.001	-0.024	0.017	-0.013	0.027	-0.016	-0.042	-0.030	-0.031	0.019
	(0.969)	(0.442)	(0.588)	(0.678)	(0.370)	(0.599)	(0.171)	(0.332)	(0.318)	(0.524)
GAPP	-0.015	0.003	-0.032	-0.022	0.022	0.035	-0.014	-0.000	0.001	0.030
	(0.632)	(0.924)	(0.299)	(0.469)	(0.466)	(0.251)	(0.639)	(0.987)	(0.977)	(0.332)
Downward	0.013	-0.006	0.001	-0.024	0.002	-0.012	-0.014	-0.013	-0.000	-0.022
	(0.677)	(0.837)	(0.972)	(0.427)	(0.943)	(0.691)	(0.648)	(0.669)	(0.994)	(0.465)

Spearman's ρ are reported. P-value in parentheses.

Table A5: OLS Regressions for Demographics and Rationality in Study 1

	Food				Risk			
	(1) CCEI	(2) HMI	(3) MPI	(4) MCI	(5) CCEI	(6) HMI	(7) MPI	(8) MCI
Female	0.017***	0.003	-0.008**	-0.002***	-0.006	-0.011	0.002	0.001
	(0.005)	(0.004)	(0.004)	0.000	(0.006)	(0.007)	(0.004)	(0.002)
Medium income	0.015***	0.010**	-0.012***	-0.001**	0.022***	0.007	-0.014***	-0.005***
	(0.005)	(0.004)	(0.004)	(0.001)	(0.007)	(0.007)	(0.004)	(0.002)
High income	0.023***	0.015***	-0.015***	-0.002***	0.046***	0.029***	-0.026***	-0.010***
	(0.006)	(0.005)	(0.005)	(0.001)	(0.007)	(0.008)	(0.004)	(0.002)
Medium education	0.001	0.004	-0.001	-0.001	0.001	0.001	-0.003	-0.000
	(0.005)	(0.004)	(0.004)	(0.000)	(0.006)	(0.007)	(0.004)	(0.002)
High education	0.000	-0.003	0.004	-0.001	0.008	0.013*	-0.007	-0.002
	(0.006)	(0.004)	(0.004)	(0.001)	(0.007)	(0.008)	(0.004)	(0.002)
Medium age	-0.008	-0.005	0.008	0.001	-0.014	-0.001	0.009	0.003
	(0.010)	(0.008)	(0.008)	(0.001)	(0.013)	(0.014)	(0.008)	(0.003)
Elder age	-0.028***	-0.012	0.018**	0.003**	-0.052***	-0.023	0.030***	0.011***
	(0.011)	(0.008)	(0.008)	(0.001)	(0.014)	(0.014)	(0.008)	(0.004)
Family size	0.011***	0.004**	-0.006***	-0.001***	-0.003	0.000	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)
N	1073	1073	1073	1073	1073	1073	1073	1073
R^2	0.080	0.027	0.046	0.082	0.088	0.044	0.080	0.063

Note: Medium (High) income indicates whether family income is 5000-10000 (more than 10000) RMB per month. Medium (High) education is high school (above high school). Medium (Elder) age refers to those born 1970-1989 (before 1969). Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: OLS Regressions for Behavioral Factors and Rationality in Study 1

	Food				Risk			
	(1) CCEI	(2) HMI	(3) MPI	(4) MCI	(5) CCEI	(6) HMI	(7) MPI	(8) MCI
Expenditure	0.013***	0.019***	-0.009**	-0.002***	0.003	-0.003	-0.001	-0.000
	(0.005)	(0.004)	(0.004)	(0.000)	(0.006)	(0.006)	(0.004)	(0.002)
Frequency	0.168***	0.055	-0.158***	-0.003	0.038	0.124*	0.011	-0.007
	(0.053)	(0.041)	(0.041)	(0.005)	(0.066)	(0.072)	(0.040)	(0.017)
IQ	-0.001	0.000	0.001	0.000	0.012***	0.004**	-0.007***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)	(0.002)	(0.001)	(0.000)
Conscientiousness	0.008***	0.004***	-0.005***	-0.000**	-0.004	0.002	0.002	0.001
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)
Extraversion	-0.001	-0.003*	0.001	0.000	-0.003	0.001	0.001	0.001
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)
Agreeableness	0.001	0.001	-0.001	-0.000	-0.003	-0.005*	-0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)
Openness	0.003	0.000	-0.002	-0.000	-0.004	-0.006*	0.003*	0.001*
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)
Neuroticism	-0.000	0.000	0.001	0.000	-0.013***	-0.009***	0.007***	0.003***
	(0.002)	(0.002)	(0.002)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1073	1073	1073	1073	1073	1073	1073	1073
R^2	0.117	0.073	0.086	0.102	0.167	0.067	0.144	0.156

Note: Control variables include gender, income, education, age dummies and family size. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Study 2

Table B1: Demographics of Subjects in Study 2

Demographic Variable	Category	Percent
Gender	*Female*	62.3%
	Male	37.7%
Age	*Born before 1969*	33.8%
	Born in 1970-1989	44.9%
	Born after 1990	21.3%
Family Monthly Income	*<RMB 5000*	20.0%
	RMB 5000-10000	29.5%
	>RMB 10000	50.5%
Education	*High school below*	38.0%
	High school	19.7%
	High school above	42.3%
Family Size	*One member*	1.3%
	Two members	15.4%
	Three members	36.7%
	Four members	32.8%
	>Four members	13.8%

Table B2: Rationality Index in Food Preference: Field

	Mean	SD	Min	Median	Max
Orginal					
CCEI	0.949	0.071	0.667	0.982	1
HMI	0.943	0.063	0.786	0.955	1
MPI	0.060	0.066	0	0.041	0.277
MCI	0.004	0.009	0	0.001	0.077
Obs: 173					
Bronars					
CCEI	0.937	0.071	0.667	0.982	1
HMI	0.943	0.063	0.682	0.955	1
MPI	0.060	0.066	0	0.041	0.277
MCI	0.004	0.009	0	0.001	0.077
Obs: 173					

Table B3: Rationality Index in Food Preference: Lab

	Mean	SD	Min	Median	Max
Original					
CCEI	0.940	0.101	0.500	0.990	1
HMI	0.920	0.100	0.409	0.955	1
MPI	0.050	0.061	0	0.028	0.274
MCI	0.017	0.055	0	0	0.517
Obs: 305					
Bronars					
CCEI	0.812	0.083	0.526	0.820	1
HMI	0.817	0.068	0.591	0.818	1
MPI	0.136	0.037	0	0.139	0.242
MCI	0.043	0.032	0	0.036	0.192
Obs: 1,000					

Table B4: Rationality Index in Risk Preference

	Mean	SD	Min	Median	Max
Original					
CCEI	0.949	0.080	0.500	0.990	1
HMI	0.908	0.089	0.591	0.909	1
MPI	0.043	0.059	0	0.020	0.443
MCI	0.009	0.027	0	0.001	0.298
Obs: 305					
Bronars					
CCEI	0.727	0.114	0.331	0.737	0.980
HMI	0.756	0.071	0.545	0.773	0.955
MPI	0.154	0.043	0.032	0.154	0.299
MCI	0.077	0.048	0.004	0.069	0.336
Obs: 1,000					

Table B5: Rationality Index in Social Preference

	Mean	SD	Min	Median	Max
Orginal					
CCEI	0.934	0.129	0.250	1	1
HMI	0.925	0.088	0.591	0.955	1
MPI	0.058	0.080	0	0.025	0.426
MCI	0.013	0.042	0	0	0.306
Obs: 305					
Bronars					
CCEI	0.677	0.133	0.352	0.679	1
HMI	0.800	0.069	0.591	0.818	1
MPI	0.216	0.061	0	0.219	0.420
MCI	0.083	0.061	0	0.070	0.465
Obs: 1,000					

Table B6: Correlations of CCEI in Study 2: Food in Field vs. Risk (Social)

	Risk										Social				
	CCEI		HMI		MPI		MCI								
	Original	*FOSD*	*Original*	*FOSD*	*Original*	*FOSD*	*Original*	*FOSD*	*EU*	*Downward*	CCEI	HMI	MPI	MCI	Downward
Food: field															
Original	0.090	0.004	0.011	0.058	0.080	0.023	0.128	0.038	-0.008	-0.060	0.061	0.024	0.002	0.033	-0.023
	(0.239)	(0.960)	(0.884)	(0.445)	(0.291)	(0.764)	(0.092)	(0.623)	(0.920)	(0.429)	(0.423)	(0.757)	(0.978)	(0.663)	(0.764)
Three years	0.095	0.000	-0.006	0.003	0.103	-0.004	0.146	0.022	0.009	-0.143	0.087	0.021	0.027	0.060	-0.023
	(0.214)	(0.996)	(0.938)	(0.965)	(0.174)	(0.961)	(0.054)	(0.772)	(0.910)	(0.359)	(0.254)	(0.780)	(0.721)	(0.434)	(0.761)
Three goods	0.092	0.069	-0.010	-0.022	0.107	0.049	0.102	0.051	-0.007	-0.035	0.058	-0.018	0.066	0.071	-0.013
	(0.225)	(0.366)	(0.896)	(0.776)	(0.158)	(0.518)	(0.182)	(0.502)	(0.928)	(0.707)	(0.444)	(0.814)	(0.389)	(0.355)	(0.866)
GAPP	0.109	0.040	0.006	0.052	0.105	0.014	0.166	0.033	0.051	-0.195	0.161	0.096	0.152	0.141	-0.120
	(0.154)	(0.602)	(0.937)	(0.494)	(0.168)	(0.853)	(0.029)	(0.662)	(0.504)	(0.118)	(0.034)	(0.207)	(0.045)	(0.063)	(0.115)
Selten score	0.020	-0.039	-0.018	0.059	0.010	0.001	0.047	-0.038	-0.011	-0.011	0.034	0.000	-0.019	0.044	-0.005
	(0.793)	(0.607)	(0.810)	(0.436)	(0.895)	(0.990)	(0.536)	(0.615)	(0.890)	(0.614)	(0.656)	(0.997)	(0.803)	(0.568)	(0.952)
Downward	-0.132	-0.037	-0.150	-0.058	0.079	0.025	0.154	0.092	-0.021	0.082	-0.092	-0.139	0.101	0.086	0.057
	(0.081)	(0.628)	(0.514)	(0.444)	(0.301)	(0.746)	(0.042)	(0.227)	(0.788)	(0.284)	(0.226)	(0.068)	(0.184)	(0.260)	(0.455)
Food: lab															
Original	0.237	0.114	0.166	0.172	0.217	0.105	0.261	0.129	0.156	-0.192	0.237	0.254	0.294	0.224	-0.192
	(0.000)	(0.046)	(0.004)	(0.003)	(0.000)	(0.066)	(0.000)	(0.025)	(0.006)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Downward	-0.235	-0.169	-0.179	-0.224	0.251	0.152	0.259	0.205	-0.192	0.202	-0.183	-0.227	0.234	0.183	0.244
	(0.000)	(0.003)	(0.002)	(0.000)	(0.000)	(0.008)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)

Table B7: Correlations of Rationality in Study 2: Social vs. Risk

	Risk									
	CCEI		HMI		MPI		MCI		*EU*	*Downward*
	Original	*FOSD*	*Original*	*FOSD*	*Original*	*FOSD*	*Original*	*FOSD*		
Social										
Original	0.396	0.347	0.426	0.331	0.433	0.384	0.409	0.353	0.218	-0.400
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Downward	-0.380	-0.382	-0.396	-0.421	0.388	0.374	0.410	0.434	-0.382	-0.233
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table B8: Correlations of Rationality of Food Preference in Study 2: Field vs. Lab

	Food: lab				
	CCEI	HMI	MPI	MCI	Downward
Food: field					
Original	0.580	0.465	0.448	0.569	-0.546
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Three years	0.520	0.472	0.400	0.550	-0.456
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Three goods	-0.190	-0.184	-0.184	-0.213	0.118
	(0.012)	(0.015)	(0.015)	(0.005)	(0.122)
GAPP	0.367	-0.341	0.288	0.236	-0.372
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Selten score	0.536	0.420	0.456	0.395	-0.407
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Downward	-0.325	-0.275	0.309	0.324	0.309
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table B9: OLS Regressions for Demographics and Rationality in Study 2

	Risk				Social			
	(1) CCEI	(2) HMI	(3) MPI	(4) MCI	(5) CCEI	(6) HMI	(7) MPI	(8) MCI
Female	-0.005	0.004	0.002	0.001	-0.020	-0.012	0.011	0.009*
	(0.009)	(0.010)	(0.007)	(0.003)	(0.015)	(0.010)	(0.009)	(0.005)
Medium income	0.016	0.017	-0.002	-0.007	0.014	0.012	-0.009	-0.005
	(0.012)	(0.014)	(0.010)	(0.004)	(0.020)	(0.014)	(0.012)	(0.007)
High income	0.046***	0.044***	-0.023**	-0.014***	0.051**	0.028**	-0.042***	-0.011*
	(0.012)	(0.014)	(0.009)	(0.004)	(0.020)	(0.014)	(0.012)	(0.007)
Medium edu	0.010	-0.010	-0.013	-0.001	0.023	0.003	-0.014	-0.011*
	(0.012)	(0.014)	(0.009)	(0.004)	(0.020)	(0.013)	(0.012)	(0.006)
High edu	0.006	0.012	-0.002	-0.002	0.023	0.013	-0.009	-0.012**
	(0.010)	(0.012)	(0.008)	(0.004)	(0.017)	(0.012)	(0.010)	(0.005)
Medium age	-0.018	-0.012	0.011	0.004	0.008	0.005	-0.011	-0.002
	(0.012)	(0.014)	(0.009)	(0.004)	(0.019)	(0.013)	(0.012)	(0.006)
Elder age	-0.049***	-0.027*	0.032***	0.010**	-0.063***	-0.039***	0.030**	0.018**
	(0.013)	(0.015)	(0.010)	(0.005)	(0.021)	(0.014)	(0.013)	(0.007)
Family size	-0.007	-0.002	0.007**	0.001	0.012	0.007	-0.002	-0.004
	(0.005)	(0.005)	(0.003)	(0.002)	(0.007)	(0.005)	(0.004)	(0.002)
N	305	305	305	305	305	305	305	305
R^2	0.168	0.090	0.118	0.085	0.155	0.120	0.166	0.127

Note: Medium (High) income indicates whether family income is 5000-10000 (more than 10000) RMB per month. Medium (High) education is high school (above high school). Medium (Elder) age refers to those born 1970-1989 (before 1969). Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B10: OLS Regressions for Demographics and Rationality in Study 2

	Food: field				Food: lab			
	(1) CCEI	(2) HMI	(3) MPI	(4) MCI	(5) CCEI	(6) HMI	(7) MPI	(8) MCI
Female	0.025**	0.021**	-0.026***	-0.004***	0.024**	0.022*	-0.013*	-0.008
	(0.011)	(0.010)	(0.010)	(0.001)	(0.012)	(0.012)	(0.007)	(0.007)
Medium income	0.024	-0.007	-0.020	0.001	0.029*	0.025	-0.018*	-0.004
	(0.015)	(0.014)	(0.014)	(0.002)	(0.016)	(0.016)	(0.010)	(0.009)
High income	0.034**	0.009	-0.029**	-0.003	0.031*	0.035**	-0.025**	0.001
	(0.014)	(0.013)	(0.013)	(0.002)	(0.016)	(0.016)	(0.010)	(0.009)
Medium edu	0.021	0.015	-0.006	-0.003	0.022	0.017	-0.016*	-0.008
	(0.015)	(0.014)	(0.014)	(0.002)	(0.016)	(0.016)	(0.010)	(0.009)
High edu	0.021*	0.011	-0.009	-0.002	0.008	-0.012	-0.003	0.002
	(0.012)	(0.011)	(0.011)	(0.002)	(0.014)	(0.014)	(0.008)	(0.008)
Medium age	-0.012	0.016	0.002	0.001	-0.011	-0.019	0.010	0.014*
	(0.014)	(0.013)	(0.013)	(0.002)	(0.016)	(0.016)	(0.009)	(0.009)
Elder age	-0.014	0.008	0.009	0.001	-0.029*	-0.039**	0.021**	0.007
	(0.015)	(0.014)	(0.014)	(0.002)	(0.017)	(0.017)	(0.010)	(0.010)
Family size	0.011**	0.006	-0.017***	-0.001	0.015**	0.012**	-0.008**	-0.008**
	(0.005)	(0.005)	(0.005)	(0.001)	(0.006)	(0.006)	(0.004)	(0.003)
N	174	174	174	174	305	305	305	305
R^2	0.146	0.068	0.160	0.107	0.084	0.083	0.095	0.048

Note: Medium (High) income indicates whether family income is 5000-10000 (more than 10000) RMB per month. Medium (High) education is high school (above high school). Medium (Elder) age refers to those born 1970-1989 (before 1969). In order measure expenditure and rationality in the field, we restrict our analyses to 174 consumers who bought both meat and vegetables for at least 7 months in the two-year period between 2019 and 2020. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B11: OLS Regressions for Behavioral Factors and Rationality in Study 2

	Risk				Social			
	(1) CCEI	(2) HMI	(3) MPI	(4) MCI	(5) CCEI	(6) HMI	(7) MPI	(8) MCI
Expenditure	-0.021	0.012	0.024	0.001	-0.042	-0.017	0.051**	-0.000
	(0.019)	(0.022)	(0.015)	(0.005)	(0.036)	(0.022)	(0.021)	(0.013)
Frequency	-0.108	-0.153	0.025	0.023	0.264	0.026	-0.146	0.006
	(0.174)	(0.203)	(0.146)	(0.046)	(0.339)	(0.203)	(0.197)	(0.119)
IQ	0.043**	0.031	-0.039**	-0.013**	0.118***	0.075***	-0.057**	-0.030**
	(0.021)	(0.024)	(0.017)	(0.005)	(0.040)	(0.024)	(0.023)	(0.014)
Conscientiousness	-0.001	-0.001	-0.001	0.001	0.002	0.007	-0.002	0.000
	(0.005)	(0.006)	(0.004)	(0.001)	(0.010)	(0.006)	(0.006)	(0.004)
Extraversion	-0.000	-0.008	0.004	-0.000	-0.013	-0.009	0.009	0.002
	(0.006)	(0.006)	(0.005)	(0.001)	(0.011)	(0.007)	(0.006)	(0.004)
Agreeableness	0.005	-0.003	-0.002	-0.002	-0.000	-0.002	0.001	-0.002
	(0.006)	(0.007)	(0.005)	(0.001)	(0.011)	(0.007)	(0.006)	(0.004)
Openness	-0.001	0.003	-0.000	0.000	-0.003	-0.008	0.005	0.002
	(0.006)	(0.007)	(0.005)	(0.001)	(0.011)	(0.007)	(0.006)	(0.004)
Neuroticism	-0.012**	-0.010	0.010**	0.002	-0.003	-0.005	0.002	0.001
	(0.006)	(0.007)	(0.005)	(0.001)	(0.011)	(0.007)	(0.006)	(0.004)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	174	174	174	174	174	174	174	174
R^2	0.259	0.181	0.237	0.190	0.186	0.264	0.223	0.158

Note: Control variables include gender, income, education, age dummies and family size. In order measure expenditure and rationality in the field, we restrict our analyses to 174 consumers who bought both meat and vegetables for at least 7 months in the two-year period between 2019 and 2020. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B12: OLS Regressions for Behavioral Factors and Rationality in Study 2

	Food: field				Food: lab			
	(1) CCEI	(2) HMI	(3) MPI	(4) MCI	(5) CCEI	(6) HMI	(7) MPI	(8) MCI
Expenditure	0.039**	0.054***	-0.051***	-0.006***	0.051*	0.074***	-0.043***	-0.008
	(0.018)	(0.016)	(0.017)	(0.002)	(0.026)	(0.025)	(0.015)	(0.013)
Frequency	-0.143	-0.394**	0.309*	0.015	0.287	0.048	-0.209	-0.158
	(0.169)	(0.151)	(0.157)	(0.022)	(0.242)	(0.235)	(0.143)	(0.121)
IQ	-0.008	0.010	0.012	0.002	-0.035	-0.025	0.015	0.024*
	(0.020)	(0.018)	(0.018)	(0.003)	(0.028)	(0.028)	(0.017)	(0.014)
Conscientiousness	0.015***	0.013***	-0.005	-0.002***	0.022***	0.021***	-0.016***	-0.007*
	(0.005)	(0.005)	(0.005)	(0.001)	(0.007)	(0.007)	(0.004)	(0.004)
Extraversion	0.002	0.006	-0.004	-0.001	-0.005	0.006	0.005	0.002
	(0.005)	(0.005)	(0.005)	(0.001)	(0.008)	(0.008)	(0.005)	(0.004)
Agreeableness	0.000	0.007	(-0.002)	-0.001	-0.008	-0.001	0.004	0.001
	(0.006)	(0.005)	(0.005)	(0.001)	(0.008)	(0.008)	(0.005)	(0.004)
Openness	0.000	-0.001	0.009*	-0.000	0.010	0.008	-0.005	-0.005
	(0.005)	(0.005)	(0.005)	(0.001)	(0.008)	(0.008)	(0.005)	(0.004)
Neuroticism	0.002	0.005	-0.001	-0.001	-0.001	-0.001	0.002	-0.003
	(0.005)	(0.005)	(0.005)	(0.001)	(0.008)	(0.008)	(0.005)	(0.004)
Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	174	174	174	174	174	174	174	174
R^2	0.218	0.210	0.237	0.198	0.206	0.221	0.251	0.123

Note: Control variables include gender, income, education, age dummies and family size. In order measure expenditure and rationality in the field, we restrict our analyses to 174 consumers who bought both meat and vegetables for at least 7 months in the two-year period between 2019 and 2020. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

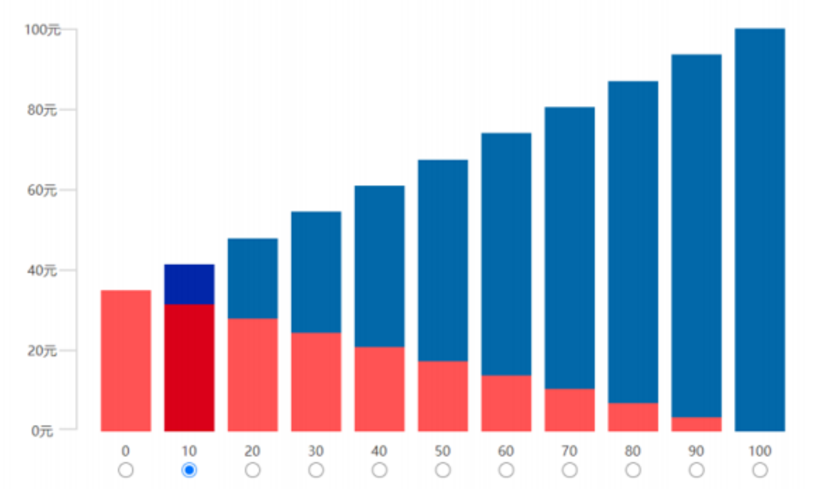
C Experiment Instructions

C.1 Risk Preference Task

As your choice may influence the bonus you receive, please ensure that you understand the following instructions and answer carefully.

- This experiment consists of 22 rounds. In every round, you have 100 points that need to be allocated between the blue membership card (Blue Card) and red membership card (Red Card).
- In the following example:
  - Blue Card: 1 point = RMB 1.00, that is, each point on the Blue Card is worth RMB 1.00
  - Red Card: 1 point = RMB 0.38, that is, each point ofn the Red Card is worth RMB 0.38
- As shown in the graph, you need to create an allocation plan. Different allocation plans influence the amount of money on your card, that is,
  - If 0 points are allocated to the Blue Card and all 100 points are allocated to the Red Card, then there will be RMB 0 on the Blue Card and RMB 38 on the Red Card;
  - If 10 points are allocated to the Blue Card and 90 points are allocated to the Red Card, then there will be RMB 10 on the Blue Card and RMB 31.5 on the Red Card;
  - ...
  - If all 100 points are allocated on the Blue Card and 0 point is allocated on the Red Card, then there will be RMB 100 on the Blue Card and RMB 0 on the Red Card.
- In each decision-making round, you have a 50% chance of getting the money on the Blue Card and another 50% chance of getting the money on the Red Card. As shown in the graph, when you choose 10 (10 points for the Blue Card and 90 points for the Red Card), you will find the following note below the bar graph "You will have a 50% chance of getting RMB 10 and another 50% chance of getting RMB 31.5". When you click on an option between 0 and 100, the corresponding amounts of money on both cards will be displayed below the bar graph.

1 token = RMB 1.00 for Blue Account and 1 token = RMB 0.35 for Red Account, which allocation would you choose?



The current choice indicates that

You will have a 50% chance of getting RMB 10.0

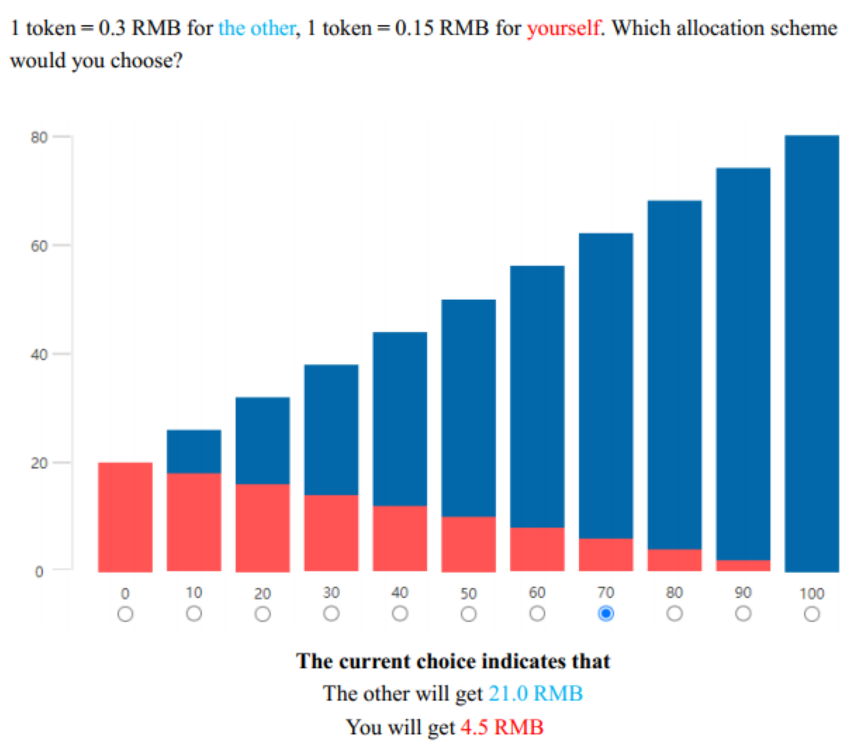
and another 50% chance of getting RMB 31.5

- Different choices will lead to different gains and risks.
  - The gains are reflected in the total length of the two-colored bars: the longer the bars are, the greater is the gain; and the shorter the bars are, the smaller is the gain.
  - Risk is reflected in the difference between the lengths of the two-colored bars: the greater is the difference in length, the greater the risk is; and the smaller is the difference in length, the smaller the risk is.
- At the end of the experiment, the computer will select one decision round randomly, where each card had an equal probability of being chosen, and the subject will be paid the amount he/she had earned in that round.

C.2 Social Preference Task

As your choice may influence the bonus you receive, please ensure that you understand the following instructions and answer carefully.

- This experiment consists of 22 rounds. In every round, you have 100 tokens, which need to be allocated between you and another supermarket customer.
- In the following example:
  - 1 token = RMB 0.3 for the other
  - 1 token = RMB 0.15 for you



- As shown in the graph, you need to create an allocation plan. Different allocation methods influence the amount of money between you and the other, that is,
  - if 0 % is allocated to the other and 100 % to yourself, the other will have RMB 0, and you will have RMB 15;
  - if 10 % is allocated to the other and 90 % for yourself, the other will have RMB 3, and you will have RMB 13.5;

  - ...
  - if 100 % is allocated to the other and 0 % tor yourself, the other will have RMB 15, and you will have RMB 0.

- As shown in the graph, when you choose 70 (30 points for the other and 70 points for yourself), you will find the following note below the bar graph "The other will get RMB 21, and you will get RMB 4.5". When you click on an option between 0 and 100, the corresponding amount of money for both will be displayed below the bar graph.
- At the end of the experiment, the computer will select one decision round randomly, and the subjects will be paid the amounts they each earned in that round.

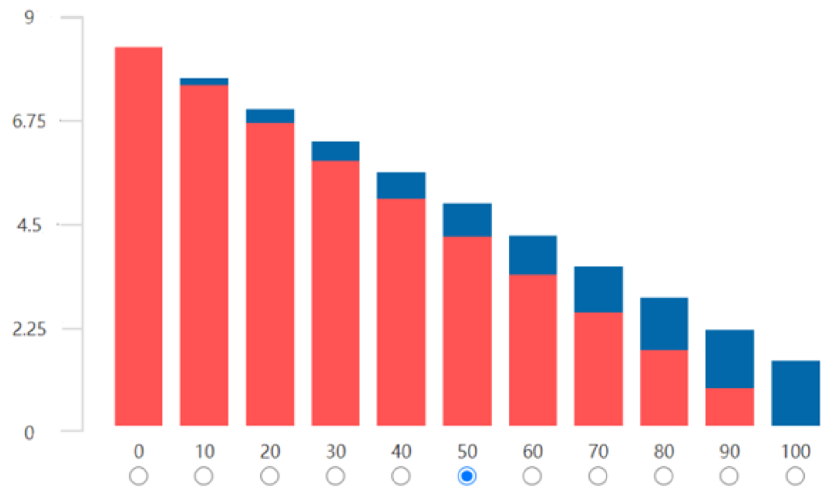
C.3 Food Preference Task

As your choice may influence the bonus you receive, please ensure that you understand the following instructions and answer carefully.

- This experiment consists of 22 rounds. In every round, you now have RMB 50 and can buy tomato or ham with these moneys.
- In the following example:
  - tomato: 6 RMB/kg
  - picnic ham: 35 RMB/kg


- As shown in the graph, you need to create an allocation plan. Different plans influence the amount of picnic ham and tomatoes you buy, that is,
  - if 0% of the expenditure is allocated to ham and 100% to tomato, you will buy 0 kg picnic ham and 8.3 kg tomatoes.
  - if 10% of the expenditure is allocated to ham and 90% to tomato, you will buy 0.14 kg ham and 7.5 kg tomatoes.
  - ...
  - if 100% of expenditure is allocated to ham and 0% to tomato, you will buy 1.43 kg ham and 0 kg tomatoes.
- As shown in the graph, when you choose 50 (RMB 25 for hams and RMB 25 for tomatoes), you will find the following note below the bar graph: "You will get 0.7 kg of ham and 4.2 kg tomatoes". When you click on an option between 0 and 100, the corresponding amounts of money allocated to each product will be displayed below the bar graph.

1 kg of hams = RMB 35, and 1 kg of tomatoes = RMB 6, which allocation would you choose?



The current choice indicates that
You will get 0.75 kg of hams
and get 4.2 kg of tomatoes

- At the end of the experiment, the computer will select one decision round randomly, and the subjects will be paid the amounts they each earned in that round.

C.4 Post-Experiment Questionnaire

- Big Five Personality Traits Test
  (7 Likert Scale) I see myself as someone who
  - is reserved
  - is generally trusting
  - tends to be lazy
  - is relaxed, handles stress well
  - has few artistic interests
  - is outgoing, sociable
  - tends to find fault with others
  - does a thorough job
  - gets nervous easily
  - has an active imagination
- Raven's IQ Test
  In each of the following questions, a part of the graph is missing from its lower right side, please find the appropriate graph to fill in the gaps. There is only one correct answer for each question.

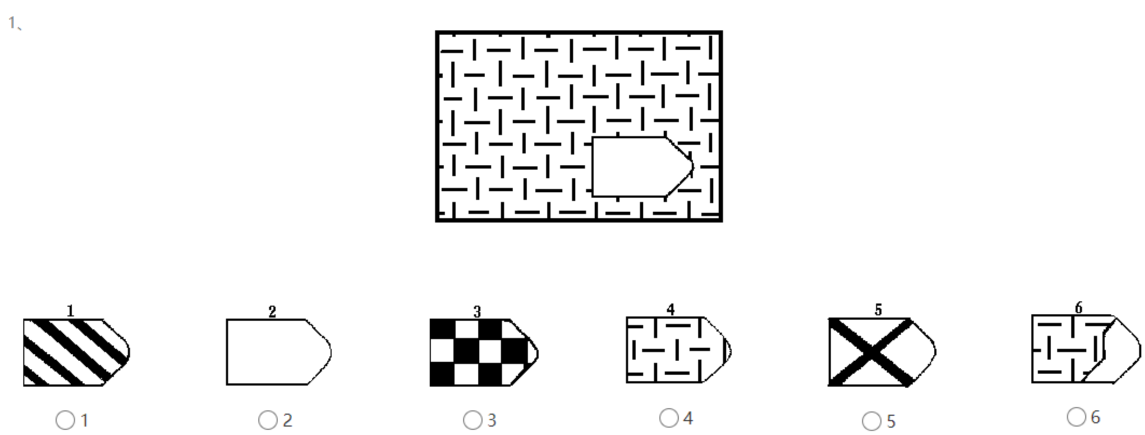


Figure C1: Sample Question of Raven's IQ Test

- Demographics
  - Gender:

- Year of Birth:
- Number of household members in your household:
- Your hukou is: Urban/Rural
- Individual monthly income:
- Household monthly income:
- Your education level: