Discussion Paper No. 1238

ISSN (Print) 0473-453X
ISSN (Online) 2435-0982

PAYING TO BE UNOBSERVABLE: EVIDENCE FROM A LYING EXPERIMENT

Te Bao
John Duffy
Nobuyuki Hanaki

Thirdly Revised September 2026
Secondly Revised February 2025
Revised September 2024
April 2024

The Institute of Social and Economic Research
The University of Osaka
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Paying to Be Unobservable: Evidence from a Lying Experiment*

Te Bao† John Duffy‡ Nobuyuki Hanaki§

September 5, 2026

Abstract

In the digital age, privacy in economic activities is increasingly threatened. In considering policies to address this threat, it is useful to consider what value, if any, people attach to being unobservable—that is, to keeping their economic activities hidden from others. We assess individuals' willingness to pay (WTP) to be unobservable in a simple coin flipping task that has been used to study lying behavior. We find that more than 90% of student participants submitted a positive WTP with the mean WTP close to 40% of the expected gain from concealment. This magnitude is replicated almost exactly in an additional experiment conducted three years later. The additional experiment separates two reasons why buyers of unobservability misreport more: participants who value unobservability more are also those who would misreport more when unobserved (selection), but having paid for it does not by itself "license" further misreporting. We also find that the observability effect reported in previous studies emerges only when the unobserved

*This is a revised version of the paper previously circulated as "Paying to avoid the spotlight." The experimental protocol governing data collection was approved by the IRB of the Institute of Social and Economic Research (ISER), The University of Osaka (Nos. 20230201, 20230803, R2026006). The experiment is pre-registered at AsPredicted.org (#147410 for the CHOICE treatment (`https://aspredicted.org/sbmw-zwqt.pdf`), #144318 for OWN and VIRTUAL treatments (`https://aspredicted.org/cbvx-r8rf.pdf`), and #297759 for the additional experiment (`https://aspredicted.org/4kz2hj.pdf`)). Lijia Wei and his assistants at Wuhan University have kindly conducted experiments in Wuhan for us. Funding was provided by (a) the Joint Usage/Research Center, ISER, The University of Osaka, (b) Grants-in-aid for Scientific Research Nos. 18K19954, 20H05631, 23H00055, and 25K24681 from the Japan Society for the Promotion of Science, Tier 1 Grant from MOE of Singapore (RG121/23), Seed Funding from NTU CCTF (026387-00004), the RIE2025 Industry Alignment Fund – Industry Collaboration Projects (IAF-ICP) (Award I2301E0026), administered by A*STAR, the Alibaba Group and NTU Singapore through Alibaba-NTU Global e-Sustainability CorpLab (ANGEL), and (c) School of Social Science, UC Irvine. We benefited from helpful comments by Shu-Heng Chen, Tony He, Juergen Huber, Lin Jing, Michael Kirchler, Marie Claire Villeval, and participants in the 2024 ESA Asia Pacific Regional Conference, the 2024 iSee workshop at NYU-Abu Dhabi, the 2024 Society for Experimental Finance Conference in Stavanger, the 2025 Innsbruck Summit on (Unethical) Behavior, WEHIA 2025 at XJTLU, Suzhou and seminars at Erasmus University Rotterdam, University of Graz and Fudan University. The authors have used Anthropic Claude and OpenAI ChatGPT for proofreading the manuscript.

†Nanyang Technological University. E-mail: baote@ntu.edu.sg

‡University of California, Irvine, and Institute of Social and Economic Research, The University of Osaka. E-mail: duffy@uci.edu

§Corresponding author. Institute of Social and Economic Research, The University of Osaka, and University of Limassol. E-mail: nobuyuki.hanaki@iser.osaka-u.ac.jp

alternative is salient: participants who were never told that an unverifiable option existed report similarly whether or not their reports can be verified.

Keywords: Privacy, Monitoring, Observabilitly, Surveillance, Willingness to pay, Dishonesty, Lying, Image concerns, Social norms

JEL Codes: C91, D83, D91

1 Introduction

In our increasingly digital world, the ability to maintain privacy in economic matters is under significant threat (Acquisti et al., 2015, 2016). The majority of financial transactions, including those conducted using debit and credit cards, online banking, and mobile payment platforms, are now conducted electronically. These methods can all be easily surveilled, with transaction records accessible through court orders in some jurisdictions, or directly monitored by governments in other settings. Consequently, such transactions often lack any privacy protections. The rise of cryptocurrencies partially reflects a pursuit of transactional privacy and independence from the volatility of sovereign currencies (Herskind et al., 2020). In response, several central banks, initially China, are developing central bank digital currencies (CBDCs) to sustain fiat currency demand and thus enable or preserve various monetary policies. However, CBDC transactions that are linked to electronic ledgers will inevitably lead to a further erosion of privacy (Ahnert et al., 2022; Wang, 2023).

There is a large literature on internet privacy that addresses topics such as the provision of personal information during online purchases (Beresford et al., 2012; Jentzsch et al., 2012; Preibusch et al., 2013; Tsai et al., 2011), and methods to hide various types of information such as browsing history, contact information, location, and text on smartphone apps (Savage and Waldman, 2015; Skatova et al., 2023).

While such concerns about controlling the disclosure of personal information are certainly important, they are not the specific focus of our study. Instead, we are interested in privacy surrounding *economic transactions*—specifically, the desire for economic activities to be unobservable, as is possible with cash transactions.

People may seek such unobservability for several reasons. It can be an expression of agency—control over the disclosure of one's own information (Van Aaken et al., 2014; Oshana, 2016), and it can shield legal but stigmatized activities, such as mental health treatment or bankruptcy, whose exposure would damage a person's social standing and thus their economic opportunities (Roessler and Mokrosinska, 2013). It can also shield illicit or dishonest conduct from reputational or legal consequences, and this motive is the one that is most germane to our study. These mixed motives are what make privacy contentious as policy: the same unobservability that protects autonomy also protects misconduct, and any measure of the value of unobservability in such settings will reflect both.

To obtain an incentive-compatible measure of the value of unobservability, we use a laboratory experiment that adapts the truth-telling task of Fischbacher and Föllmi-Heusi (2013).[1] Participants flip a fair coin ten times and report the outcome; each head reported earns money, so the most profitable report is ten heads, and there is no penalty for misreporting.[2] What we add is a choice of coin. Participants may use either a virtual coin provided by our software, whose realizations we observe and can compare with their reports, or their own physical coin, whose outcomes we cannot observe. The virtual coin is free, while using one's own coin requires a payment. We elicit each participant's willingness to pay to use their own coin using the Becker–DeGroot–Marschak (BDM) mechanism (Becker et al., 1964).

Our design provides a highly stylized version of the tradeoff between verifiability and privacy in economic transactions. Participants report a payoff-relevant outcome to

[1]There is a large experimental literature on lying that explores behavior in tasks like coin-flipping and dice-rolling, where participants often misreport outcomes to their advantage, especially under conditions of anonymity (Fries et al., 2021; Kajackaite and Gneezy, 2017; Gneezy et al., 2018). However, the extent of dishonesty appears to be limited by concerns over self-image and social perception. Gächter and Schulz (2016) found that in countries with prevalent rule violations, participants were more likely to report favorable outcomes. For a comprehensive review of this literature, see the meta-analysis by Abeler et al. (2019).

[2]In contrast to most research exploring the value of privacy (Prince and Wallsten, 2022; Savage and Waldman, 2015; Skatova et al., 2023; Tsai et al., 2011), we elicit people's WTP for privacy in an incentive-compatible manner. Indeed, Benndorf and Normann (2018) show a large "hypothetical bias" in people's willingness to disclose their personal information: namely, while five of six respondents in a non-incentivized survey refused to disclose their personal information, only one of six participants did the same in an incentivized experiment. Our incentivized experiment does not suffer from a similar bias.

experimenters using either a free channel whose record the experimenters can verify, or a costly channel whose record they cannot. This parallels the choice between traceable payment methods such as credit cards or CBDC and untraceable methods such as cash, with the fee for using one's own coin representing the price of the privacy-preserving option. What our experiment measures is therefore not the value of privacy in general, but something narrower and more precise: the value of unobservability in a single transaction with a known monetary stake.

We ask three questions. First, is the WTP for unobservability positive in this setting, and if so, how large is it relative to the material gain that concealment makes available? Second, does the availability of a purchasable private option change reporting behavior, relative to environments in which observability is imposed exogenously? Third, *why* do those who buy unobservability misreport more? Two mechanisms are possible. First, participants who place a higher value on unobservability may simply be those who intended to misreport more in the first place (a *selection* effect). Second, paying for unobservability may by itself justify misreporting that would not otherwise have occurred (a *licensing* effect). We distinguish between these mechanisms in Experiment 2. All participants first submit their WTP to use their own coin, but the BDM mechanism is implemented for only a randomly selected half of participants. For the remaining half, the coin is assigned randomly and no fee is charged.

Participants in our experiments have a large WTP to conceal their misreporting. Mean WTP is 188.09 points (s.d. 143.68), close to 40% of the expected gain from maximum misreporting (500 points), with 92.8% of participants submitting a positive value. This result is replicated almost exactly in Experiment 2, conducted three years later with the same subject pools: mean WTP of 188.35 points (s.d. 150.71), with 95.8% of participants submitting a positive value.

Our finding also suggests that the observability effect (Fries et al., 2021; Kajackaite and Gneezy, 2017; Gneezy et al., 2018) may depend on its salience. In Experiment 1, we found no significant difference in reporting behavior between participants assigned the virtual coin and those assigned their own coin, when neither group was informed that an alternative type of coin existed. In Experiment 2, by contrast, all participants were informed about the two types of coins and asked to submit their WTP to use their own coin. In this setting, we find a significant difference in reporting behavior: participants randomly assigned their own coin reported roughly one more head, on average, than those randomly assigned the virtual coin.

Finally, those participants who paid to use their own coin reported more heads than those who were randomly assigned to use their own coin. Our Experiment 2 reveals that this was primarily due to the selection effect, and not due to the licensing effect. Participants who submitted a higher WTP report more heads when they happen to be given their own coin, but not when they happen to be given the virtual coin. However, conditional on their WTP, those who paid to use their own coin do not report significantly more heads than those who were randomly assigned to use their own coin without any fee.

To the best of our knowledge, this is the first paper to measure the value of unobservability when it can conceal misreporting in a certain activity, here, in the lying game. We also distinguish between two reasons why participants who pay for unobservability misreport more than those who are assigned an unobservable option: selection and licensing. We find that the difference is driven primarily by selection rather than licensing. Furthermore, our findings indicate the effect of observability on lying behavior may depend on its salience.

Finally, we conducted the experiment in three university subject pools—in Osaka (Japan), Wuhan (China) and Irvine, California (United States)—to assess whether WTP for unobservability is specific to a particular subject pool. It is not: WTP is positive and economically large in all three. While significant differences in WTP across locations exist, they are not stable across two waves. We report the location differences briefly in Section 3.1 and relegate the location-by-location analysis to Online Appendix E. The main analyses of the paper are based on pooled data.

The remainder of the paper is organized as follows. Section 2 describes the design, the conceptual framework and the implementation of the two experiments. Section 3 reports the results. Section 4 offers a concluding discussion.

2 Experimental Design

In our experiment, as in Cohn et al. (2014), participants flip a fair coin 10 times and report the number of heads and tails, which together must add up to 10. Participants earn 100 points for each head reported. Thus, a report of 10 heads yields the maximum payoff of 1,000 points. At the conclusion of the experiment, these points are converted into local currency using a pre-specified exchange rate that adjusts for purchasing power differences across countries, ensuring that all participants face roughly equivalent monetary incentives.[3]

There is no explicit penalty for misreporting the number of heads and tails in any treatment of our study. Therefore, the predicted, profit-maximizing behavior by *homo economicus* participants is to always report 10 heads. A meta-study of such truth-telling experiments by Abeler et al. (2019), however, shows that most participants do not lie to such an extreme degree; rather, they partially over-report the number of heads relative to the expected number. What we add to this task is an elicitation of willingness to pay (WTP) to use one's own coin rather than the free virtual coin.

2.1 Experiment 1

There are three treatments: CHOICE, OWN, and VIRTUAL. In CHOICE, which is our main treatment, subjects can complete the task of flipping a coin 10 times either by using a virtual coin provided on our experimental software platform at no cost or by using their *own* coin for a fee. In OWN and VIRTUAL, which are control treatments, participants are not given a choice of coins and are not aware of any alternative coin option. They are simply told to use the coin (the virtual coin in VIRTUAL and their own coin in OWN) without any fee. The literature (see, e.g., Abeler et al., 2019; Fries et al., 2021; Kajackaite and Gneezy, 2017; Gneezy et al., 2018) suggests that participants tend to lie

[3]Specifically, 100 points equated to 100 JPY in Japan, 1.00 USD in the United States, and 4 RMB in China. These conversion rates were purchasing power equivalent at the time of the first experimental study.

more when their behavior is more anonymous. Thus, we expect participants to report a larger number of heads in OWN than in VIRTUAL.[4]

Subjects are told that if they use the virtual coin, the experimenters can check the realized outcomes of the virtual coin flip later. By contrast, if they use their own coin, it is not possible for the experimenter to observe the outcome of those coin flips. Subjects are also told that regardless of the coin they use, the experimenters rely only on their own self-report of the outcome of the coin flip—the number of heads and tails that *they* report to the experimenter—to determine their payment.

In CHOICE, prior to the coin flip task, we use the BDM mechanism to elicit subjects' WTP for the right to use their own coin. Specifically, subjects submit their WTP^i in 10-point increments, $\{0, 10, 20, \ldots, 490, 500\}$.[5] Once participants submit their WTP^i, the computer randomly draws a price (in points) $p^i \in \{10, 20, \ldots, 490, 500\}$ for each participant. If $p^i \leq WTP^i$, subject i pays p^i (the fee) out of their experimental earnings and uses their own coin; otherwise, the virtual coin is used at no cost. Subjects who do not want to use their own coin can simply state that their WTP is zero, thereby ensuring that they use the virtual coin, and this possibility is carefully explained to them.[6]

2.2 Conceptual framework

Given that using the virtual coin is free, that payoffs are determined solely by participants' self-reported number of heads, and that there is no penalty for misreporting, a purely rational economic agent (i.e., *homo economicus*) who has no image or privacy concerns would maximize her payoff by reporting 10 heads and expressing a WTP of zero, thereby

[4] García-Gallego et al. (2020) also find that people behave differently in the corruption game when they are observed by an audience compared to when there is no audience.

[5] Note that the upper bound of 500 points in the WTP elicitation is the expected gain (in points) from using one's own coin and reporting 10 as the number of heads.

[6] BDM is known to confuse participants (Cason and Plott, 2014). We therefore walked them through the decision as an ascending sequence of yes/no questions—10 points, then 20, and so on, until the answer was no—which is equivalent to a multiple price list (MPL), and we included two BDM comprehension questions in the quiz. In Experiment 2, the quiz also included a question asking why it is optimal to submit one's true WTP. Brebner and Sonnemans (2018) find BDM to be no more confusing than MPL. Payoffs are denominated in points rather than local currency so that the instructions are identical across locations, with conversion at purchasing power parity. Pilot sessions in Osaka, conducted in local currency, produced a WTP distribution not significantly different from the CHOICE treatment in Osaka (Online Appendix F).

always using the virtual coin.

We thus interpret the elicited WTP as reflecting participants' image and privacy concerns regarding the use of the virtual coin versus their own. To formalize this logic, we adopt the framework proposed by Abeler et al. (2019) and consider (1) material gain, (2) self-image concerns, and (3) privacy concerns as three determinants of reporting behavior in this experiment. Following this framework, the utility of reporting the number of heads, H, when the actual realization is R, $u^i(H|R)$, can be written as

$$\begin{aligned} u^i_{vc}(H|R) &= \pi^i(H) - c^i(H-R) - \gamma^i(H-R) && \text{if using the virtual coin,} \\ u^i_{oc}(H|R) &= \pi^i(H) - c^i(H-R) && \text{if using one's own coin.} \end{aligned}$$

Here, $\pi^i(H)$ is the monetary gain, $c^i(H-R)$ is the cost associated with self-image concerns, and $\gamma^i(H-R)$ is the privacy concern of reporting H heads when the actual realized number of heads was R. Intuitively, the first order derivatives of $c^i(H-R)$ and $\gamma^i(H-R)$ with respect to $H-R$ should be non-negative: that is, the cost of self-image and privacy concerns should be greater when the participant deviates from honest reporting to a larger degree.[7]

When choosing a WTP to use their own coin instead of using the virtual coin for free, each subject i compares the expected maximized utility of using the virtual coin with that of using their own coin. In addition, there could be a social-image concern associated with submitting a positive WTP, as it is observed by the experimenter and can be interpreted as a willingness to misreport. Let $\sigma^i(WTP)$ represent such a concern. Intuitively, its first derivative should be non-negative. Therefore, the submitted WTP should be equivalent to the difference between the monetary value of the two maximized expected utilities, which is primarily driven by the privacy concerns in using the virtual coin ($\gamma^i(H-R)$), net of the social image concern in submitting the chosen amount of

[7] Grossman (2015) also studies how self-image and social-image motives may lead to different actions and outcomes in behavioral games. Specifically, he varies the possibility that the dictator's choice is implemented in a dictator game. He suggests that when outcomes are publicly visible, individuals primarily motivated by social image will not alter their contributions depending on the probability of their choices being implemented. In contrast, those driven by self-image concerns are likely to increase their contributions as the probability that the dictator's decision will be implemented rises.

WTP ($\sigma^i(WTP)$). Because we do not separately measure the social-image concern with respect to the WTP, our elicited WTP provides a lower bound on the monetary value of the privacy concern in our experiment.[8]

2.3 Experiment 2

In Experiment 2, all participants first submit a WTP to use their own coin. Then, for a randomly chosen half of participants, the coin they use is determined based on the BDM method as in the CHOICE treatment (henceforth, the "WTP-based" procedure).[9] For the remaining half, however, the coin they use is determined randomly as in the OWN and VIRTUAL treatments (henceforth, the "Random-assignment" procedure).[10] In the WTP-based procedure, participants must pay the randomly chosen price to use their own coin, while in the Random-assignment procedure, it can be used for free. The use of the virtual coin is free in both procedures. Participants are informed about both the WTP-based and Random-assignment procedures, but they are not told to which procedure they are assigned when submitting their WTP to use their own coin.

2.4 Questionnaire

In all the sessions, after participants reported the number of heads and tails (which must sum up to 10), they completed a questionnaire regarding their age, gender, willingness to take risks, and beliefs about others' WTP and reporting behavior. Participants were not incentivized for accuracy of these beliefs. In addition, they answered questions related to their views on ethical behavior and the government's authority to monitor economic

[8]In addition to privacy and social-image concerns, a participant may also submit a positive WTP to use their own coin because of the illusion of control (Langer, 1975). Another possibility is that participants submit a positive WTP value because they do not believe the virtual coin is a fair coin, even though the coin is in fact fair, and we inform subjects of this fact in the instructions. Although we cannot eliminate this possibility for submitting a positive WTP, the fact that our experiments were conducted in subject pools with a no-deception rule, of which participants are aware, should reduce the importance of this concern.

[9]This is referred to as the "WTP-price procedure" in the instructions reproduced in Online Appendix A.

[10]Specifically, we randomly formed groups of four participants within each session whenever possible. Within each such group, two participants were assigned to the WTP-based procedure and two to the Random-assignment procedure; of the latter two, one was assigned their own coin and the other the virtual coin. Procedures were randomly assigned to any remaining participants.

activities.[11] We will later use the answers to these questions as potential covariates in our regression analysis. The English instructions were translated into Japanese and Chinese by our research assistants. We then asked different people to translate the instructions back into English to ensure consistency of meaning.[12]

2.5 Implementation

The experiments were conducted online between October and December 2023 (Experiment 1) and between July and August 2026 (Experiment 2) in Osaka (Japan), Irvine, California (United States), and Wuhan (China).

In each location, we used the Zoom software to coordinate activity. Subjects arrived via the Zoom waiting room. One by one, we privately welcomed them and checked (via video) that they had brought their own coin to the study as they had been instructed to do (except for the treatment with only a virtual coin). They were then given a numerical ID to be used during the experiment to maintain their anonymity. Once these tasks were completed, they were sent back to the waiting room, where they waited until the start of the experiment.

Once all subjects had been individually welcomed, they were brought back to the main

[11]Specifically, in addition to age and gender, the following questions were asked: (1) How many of 20 randomly chosen participants in the experiment have reported the outcome of coin flips truthfully? (2) How many have reported a $WTP > 0$? (Not in OWN and VIRTUAL) (3) In general, how willing are you to take risks? Please indicate on the scale below from 0 to 10, where 0 means you are "completely unwilling to take risks" and a 10 means you are "very willing to take risks." (4) What do you think is the purpose of the experiment? (5) Please indicate whether you think the following actions can be always justified, never be justified or something in between using the given scale. (1: Never justifiable ... 10: Always justifiable). (a) Claiming government benefits to which you are not entitled. (b) Cheating on taxes if you have a chance. (c) Telling the truth when it is costly for you to do so. (6) Do you think your country's government should or should not have the right to do the following (1: Definitely should have the right. 2: Probably should have the right. 3: Probably should not have the right. 4: Definitely should not have the right.): (a) Keep people under video surveillance in public areas. (b) Monitor all e-mails and any other information exchanged on the internet. (c) Access to people's bank account balances and their history of payments. Question (3) is from the Global Preference Survey (Falk et al., 2018). Question (5) is related to ethics, and question (6) is related to the government's rights. Questions (5-a), (5-b), (6-a), and (6-b) are from the World Values Survey Wave 7 (Haerpfer et al., 2022). From question (5) we construct a variable called "Ethics." Namely, Ethics = ((11 - claiming benefit) + (11 - cheating tax) + telling truth)/3. A higher value of the Ethics variable indicates that a participant considers unethical behavior to be less justifiable. From question (6) we construct a variable called "Government's right." Namely, Government's right = ((5-video surveillance) + (5-internet) + (5-bank account))/3. A higher value of the Government's right variable indicates that a participant agrees to a larger extent that the government has the right to monitor people.

[12]The English instructions are available in Online Appendix A; the Chinese and Japanese translations are available upon request.

Zoom room, where the experimenter gave general instructions and sent each subject a link to the experimental platform. After clicking on the link, subjects read through the instructions for the experiment online at their own pace and then completed a comprehension quiz. Once they had answered all of the quiz questions correctly, the experiment started. In the CHOICE treatment with a WTP elicitation, participants first underwent two practice rounds to learn how to submit their WTP and its impact on the type of coin—own or virtual—they would use. One practice round featured a high price realization and the other a low price realization. Note that these practice rounds did not involve flipping a coin. Participants submitted their WTP and received feedback about the chosen price and whether they would have used their own coin or the virtual coin if these were not practice rounds. During the experiment, subjects had their cameras and microphones turned off.[13]

A total of 719 students participated: 360 in Experiment 1 and 359 in Experiment 2, with approximately 120 per location in each experiment. All participants were recruited from the experimental subject pools of the local universities. Each participant took part in only one session and those who participated in Experiment 1 were excluded from Experiment 2.[14]

The experiment lasted about 25 minutes on average, including the post-experimental questionnaire. In Experiment 1 (2), subjects, on average, earned 13.30 (13.14) USD, 1130 (1140) JPY, and 50 (47.3) RMB, including 7 USD, 500 JPY, and 20 RMB show-up fees in Irvine, Osaka, and Wuhan, respectively.[15]

[13]The experiment was programmed using oTree (Chen et al., 2016). The experiments were conducted by one of the authors in Osaka, another in Irvine, and by local researchers in Wuhan who are not co-authors of this paper. For Experiment 1, there were 2 sessions of CHOICE, one session each for OWN and VIRTUAL in Osaka and Wuhan. In Irvine, there were 5 sessions for CHOICE and VIRTUAL, and 2 sessions for OWN. For Experiment 2, there were 4 sessions in Osaka, 2 sessions in Wuhan, and 11 sessions in Irvine.

[14]In Experiment 2 in Osaka, where we had 122 participants, 3 of them could not go beyond the practice WTP submission stage due to a technical problem. We have dropped them from the sample.

[15]While performance-based payments were equalized based on purchasing power parity, we respected the standard show-up fees used in each location so as to attract participants who normally participate in experiments conducted in each location.

3 Results

Among the 360 participants in Experiment 1, 183 are in CHOICE, 90 are in OWN, and 87 are in VIRTUAL. Among the 359 participants in Experiment 2, 174 are in WTP-based and 185 are in Random-assignment; of the latter, 96 were assigned to use their own coin and 89 to use the virtual coin.

Descriptive statistics on participants' characteristics by location and treatment are reported in Tables B.1 and B.2 of Online Appendix B for Experiments 1 and 2, respectively, along with p-values from Kruskal–Wallis (KW) tests for differences across the three locations. Table B.3 reports treatment- and experiment-level averages pooling data from all three locations, together with p-values from Kruskal-Wallis (KW) tests for differences across treatments.

While there are some notable differences in participants' characteristics across the three locations depending on the experiment, there are no significant differences across treatments. We therefore control for these individual characteristics as well as location and experimental wave, in the analyses below.

3.1 Willingness to pay to use one's own coin

We begin with the main variable of interest: subjects' WTP to use their own coin instead of the virtual coin. We then present the reporting behavior of participants conditional on the coin they actually used.

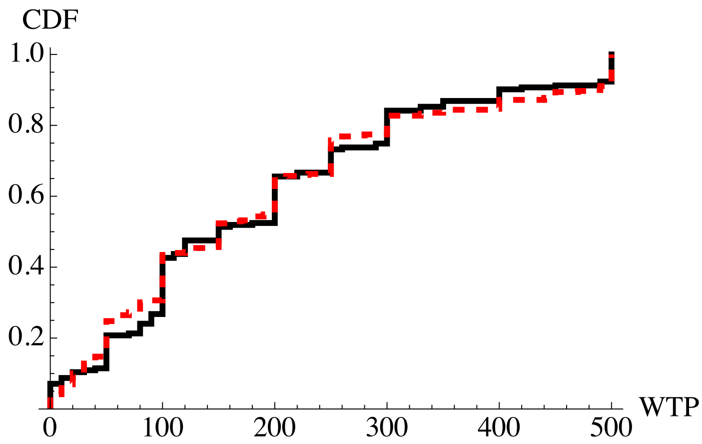
Figure 1 shows the CDFs of subjects' WTP to use their own coin instead of using the virtual coin for free in Experiments 1 (Solid Black) and 2 (Dashed Red). The mean WTPs (std. devs.) are 188.09 points (143.68) and 188.35 (150.71) in Experiments 1 and 2, respectively. There is no significant difference between the two experiments (p=0.719, Mann–Whitney (MW) test, two-tailed).

WTP varies considerably across participants. In both Experiments 1 and 2, WTP spans the entire elicitation range: there are participants who submit the minimum WTP (0) as well as those who submit the maximum (500). In the regression analyses below,

Table 1: Willingness to pay (WTP) to use their own coin: Results of ordinary least squares regressions.

	Exp 1		Exp 2		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Wuhan	−65.722***	−61.625**	39.282*	44.855**	−65.722***	−52.264**
	(25.051)	(25.367)	(20.502)	(22.237)	(24.983)	(23.335)
Irvine	−48.556*	−39.090	−2.718	24.581	−48.556*	−27.188
	(26.754)	(26.947)	(18.172)	(20.472)	(26.682)	(25.543)
Exp 2					−49.421**	−46.891**
					(23.355)	(21.862)
Exp 2 × Wuhan					105.004***	90.429***
					(32.336)	(29.656)
Exp 2 × Irvine					45.838	46.455
					(32.296)	(30.665)
Female		−5.315		−25.808		−18.063
		(19.919)		(16.155)		(12.540)
Other		52.420		−59.384		−8.234
		(45.773)		(60.855)		(41.666)
Age		0.570		−0.458		0.748
		(3.343)		(3.539)		(2.502)
Risk		16.611***		6.386*		10.260***
		(4.449)		(3.594)		(2.773)
Ethics		−3.539		2.438		−0.956
		(6.032)		(5.976)		(4.248)
Gov. right		17.714		16.117		18.519**
		(15.001)		(11.593)		(9.024)
No. Truthful		−5.563***		−3.708**		−4.495***
		(1.721)		(1.577)		(1.173)
No. WTP>0		6.472***		9.078***		8.136***
		(1.909)		(1.481)		(1.151)
Constant	225.556***	73.948	176.134***	6.827	225.556***	48.243
	(18.964)	(96.078)	(13.683)	(83.796)	(18.913)	(64.434)
Adjusted R^2	0.028	0.210	0.011	0.150	0.014	0.167
N	183	180	359	359	542	539
H_0: Wuhan = Irvine	0.4928	0.3882	0.0310	0.3570		
H_0: location × Exp 2 = 0^a					0.0051	0.0065
H_0: characteristics, beliefs × Exp 2 = 0^a						0.4794

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Exp 1 is the CHOICE treatment; Exp 2 is the full sample. Columns (5) and (6) are saturated in experiment × location, so the location coefficients in column (5) reproduce those in column (1) exactly.
The final three rows report p-values from Wald tests of the stated null hypotheses.
a: obtained from a regression allowing the coefficients on location, individual characteristics and beliefs to differ between the two experiments (Table D.1 in Online Appendix D). Because equality of the characteristics and belief coefficients is not rejected, column (6) imposes it; location effects are left unrestricted throughout.
N is smaller in columns (2) and (6) because the Ethics measure is missing for 3 participants owing to a technical problem.



	Exp. 1	Exp. 2	*p*-value*
Mean	188.09	188.35	0.719
(s.d.)	(143.68)	(150.71)	

* p-value from Mann-Whitney test, two-tailed.

Figure 1: Cumulative Distributions of willingness to pay to use one's own coin in Experiments 1 (Solid Black) and 2 (Dashed Red).

we ask which individual characteristics account for that dispersion, whether the same characteristics matter in the two samples drawn three years apart, and how much of the variation they can explain.

Table 1 reports ordinary least squares (OLS) regressions of WTP on location dummies and individual characteristics, including demographics, risk preferences, views on ethical behavior and the government's right to monitor people, and beliefs about other participants' behavior. Columns 1–2 report results for Experiment 1, Columns 3–4 for Experiment 2, and the remaining columns pool the data from both experiments.[16] Heteroskedasticity-robust standard errors are reported in parentheses.[17]

While the average WTP is very similar in Experiments 1 and 2, there are variations across locations which are not stable across time. On the one hand, in Experiment 1, the WTP is significantly lower in Wuhan compared to Osaka (the baseline). On the other

[16]The analyses using the pooled data were not pre-registered.

[17]We report heteroskedasticity-robust rather than session-clustered standard errors for two reasons. First, there are only 20 sessions in Experiment 1, nine of which involve the CHOICE treatment, and four treatment-by-location cells contain only a single session; Experiment 2 has even fewer sessions. Cluster-robust variance estimates based on so few clusters are known to be downward-biased (Cameron et al., 2008; MacKinnon and Webb, 2017). Second, consistent with this concern, session-clustered standard errors are smaller than heteroskedasticity-robust standard errors in almost every specification we estimate (see Online Appendix G), making the latter the more conservative choice. Intra-session correlations are also very low: 0.0215, 0.0344, and 0.027 for submitted WTP in Experiment 1, Experiment 2, and the pooled data, respectively, and 0.000, 0.046, and 0.023 for reported heads.

hand, in Experiment 2, it is the opposite. Thus, we will not try to interpret the locational differences in WTP.[18]

The relationship between individual characteristics and WTP to use their own coin is stable over time. Namely, the self-reported willingness to take risks is positively correlated with WTP, while age, gender, and participants' views on ethical behavior are not. In the pooled data, we observe a positive and significant relationship between participants' views on government's right to monitor people and their WTP. That is, the more accepting participants are of their government monitoring people's activities, the higher is their WTP to use their own coin.

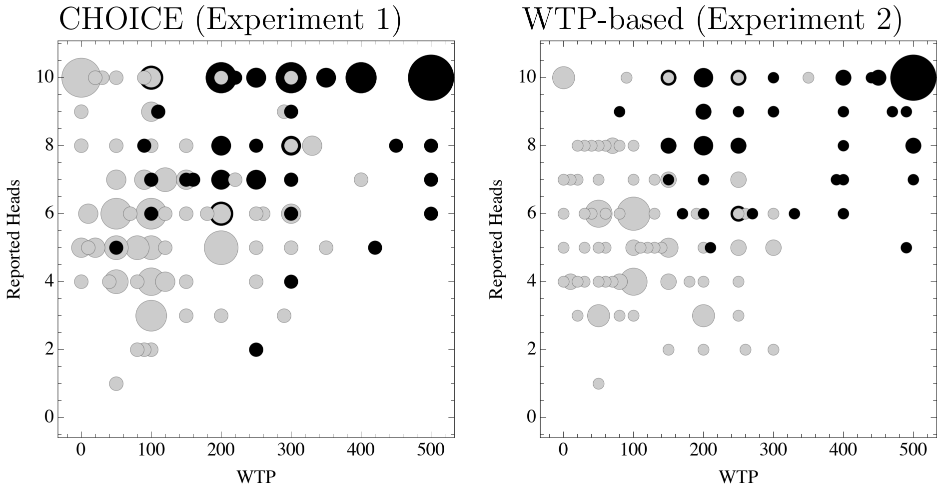
We also observe significant relationships between participants' beliefs about the behavior of others and WTP. "No. Truthful" and "No. WTP>0" are their beliefs about how many of 20 randomly selected other participants report their coin-flip outcomes truthfully and submit a positive WTP, respectively.[19]

Columns (2), (4), and (6) show that while WTP is negatively related to "No. Truthful", it is positively related to "No. WTP>0." That is, the less likely participants believe that others report truthfully, or the more likely participants believe that others submit a positive WTP to use their own coin, the higher is their own WTP.

We interpret these two correlational findings as consistent with a role for social norm considerations (Cialdini and Goldstein, 2004; Gächter and Schulz, 2016; Serdarevic, 2021) in shaping participants' *social-image concerns* about submitting a positive WTP to use their own coin, represented by σ^i(WTP) in our conceptual framework. Participants who believe that others are more likely to report truthfully may view a higher WTP as carrying a greater social-image cost, and therefore submit a lower WTP. By contrast, participants who believe that others are more likely to report a positive WTP may perceive a lower social-image cost and therefore submit a higher WTP. Although these relationships are purely correlational, they are consistent with our conceptual framework and may help explain variation in WTP across participants.

[18]See Online Appendix E for location level results.

[19]As noted above, belief elicitations are not incentivized, and are conducted at the end of the experiment. Thus, the interpretation below is subject to ex-post justification.



	Own	Virtual
N	65	118

	Own	Virtual
N	63	111

Figure 2: Submitted WTP and the reported number of heads depending on the coin used: Own coin (black); virtual coin (gray). CHOICE treatment in Experiment 1 (left) and WTP-based procedure in Experiment 2 (right). The size of the point is proportional to the number of observations for the same WTP-report combination.

Taken together, the correlates of WTP are stable across two independent samples: willingness to take risks and beliefs about others' behavior are associated with WTP in both experiments, while demographics and stated views on ethics and the government's right to monitor people are not. The replication of these relationships across independent samples provides stronger evidence than any single-sample specification alone.

3.2 Reporting behavior

Let us turn next to the relationship between participants' WTP to use their own coin and their reported number of heads when the BDM mechanism is used to determine the coin. Figure 2 uses bubble plots to show the relationship between subjects' WTP values (horizontal axis) and their reported number of heads (vertical axis), depending on the coin used (own coin in black and virtual coin in gray). The CHOICE treatment of Experiment 1 is shown in the left panel and the WTP-based procedure of Experiment 2 is shown in the right panel.

In both experiments, some subjects submitted a WTP of 0 and reported 10 heads (as *homo economicus* would do), as can be seen in the upper left corner of the two panels of

Figure 2. We also see in the upper right corner that some subjects submitted a WTP of 500 and reported 10 heads.

Let us investigate this relationship further. Table 2 shows the results of OLS regressions. On the one hand, we observe that, when their own coin is used (the default condition) there is a positive correlation between WTP and the reported number of heads except when beliefs about others' behavior are controlled for in Experiment 1 (Column (2)). On the other hand, when the virtual coin is used (WTP $\times$ Virtual coin) there is no significant correlation between WTP and the reported number of heads (see the row marked WTP + WTP$\times$ Virtual=0 of Table 2 that reports p-values for this test). This suggests that, in line with our conceptual framework, those with a higher WTP to use their own coin are more likely to try to benefit from doing so.

The estimated coefficients on "No. Truthful" are negative and significant, indicating that participants who believe others are more likely to report truthfully report fewer heads. As noted above, because these beliefs are elicited without incentives and only after the experiment, this evidence is purely correlational. Nevertheless, it is consistent with a role for social norms in participants' reporting behavior.

Online Appendix C shows that among participants in the CHOICE treatment of Experiment 1 and the WTP-based procedure of Experiment 2, those who used their own coin earned significantly more than those who used the virtual coin. Thus, at least in terms of realized earnings, participants who paid for the right to use their own coin can be viewed as having made a "rational" choice.

3.3 Effect of observability

The two panels in Figure 3 compare the CDFs of the reported number of heads when the coin is randomly assigned, allowing us to examine the effect of observability. Panel (A) compares the OWN (Solid Black) and VIRTUAL (Dashed Gray) treatments in Experiment 1, where participants were not aware that an alternative type of coin existed. Panel (B) compares participants randomly assigned to use their own coin (Solid Black) with those randomly assigned to use the virtual coin (Dashed Gray) under the Random-assignment

Table 2: Reported number of heads and willingness to pay: OLS regression results

	Exp 1 (CHOICE)		Exp 2 (WTP-based)		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Wuhan	1.014***	0.981**	-0.046	-0.271	1.030***	0.877**
	(0.374)	(0.381)	(0.356)	(0.404)	(0.370)	(0.348)
Irvine	-0.073	-0.090	-0.834**	-0.561	-0.056	-0.053
	(0.380)	(0.403)	(0.355)	(0.394)	(0.373)	(0.372)
Exp2					0.172	0.114
					(0.357)	(0.326)
Wuhan×Exp2					-1.095**	-1.058**
					(0.512)	(0.474)
Irvine×Exp2					-0.741	-0.489
					(0.506)	(0.485)
WTP	0.005***	0.002	0.003**	0.003***	0.004***	0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Virtual coin	-0.943	-1.519**	-1.748***	-1.035	-1.270***	-1.195***
	(0.638)	(0.626)	(0.607)	(0.637)	(0.454)	(0.432)
WTP × Virtual coin	-0.006**	-0.001	-0.006**	-0.008**	-0.006***	-0.005**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
No. Truthful		-0.176***		-0.109***		-0.144***
		(0.028)		(0.031)		(0.020)
No. WTP>0		-0.035		0.001		-0.023
		(0.035)		(0.033)		(0.024)
Constant	6.891***	8.448***	7.939***	8.945***	7.227***	8.591***
	(0.514)	(1.624)	(0.565)	(2.006)	(0.400)	(1.223)
Individual characteristics	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.266	0.384	0.432	0.470	0.349	0.431
N	183	180	174	174	357	354
p: WTP + WTP×Virtual = 0	0.6640	0.6614	0.2595	0.6440	0.2841	0.3976
p: location × wave = 0	–	–	–	–	0.1094	0.0748
p: WTP terms × wave = 0	–	–	–	–	0.4768	0.2789
p: characteristics × wave = 0	–	–	–	–	–	0.6400

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Individual characteristics are gender, age, willingness to take risks, ethics and views on government monitoring; none is significant in any specification and coefficients are reported in Table D.2 in Online Appendix D.
The last three rows report joint tests of whether coefficients differ between the two experiments, obtained from a regression in which every coefficient is allowed to differ by wave. “WTP terms” are WTP, the virtual-coin indicator and their interaction.
N is smaller for Columns (2) and (6) than for Columns (1) and (5) because we are missing Ethics for 3 participants due to technical problems.

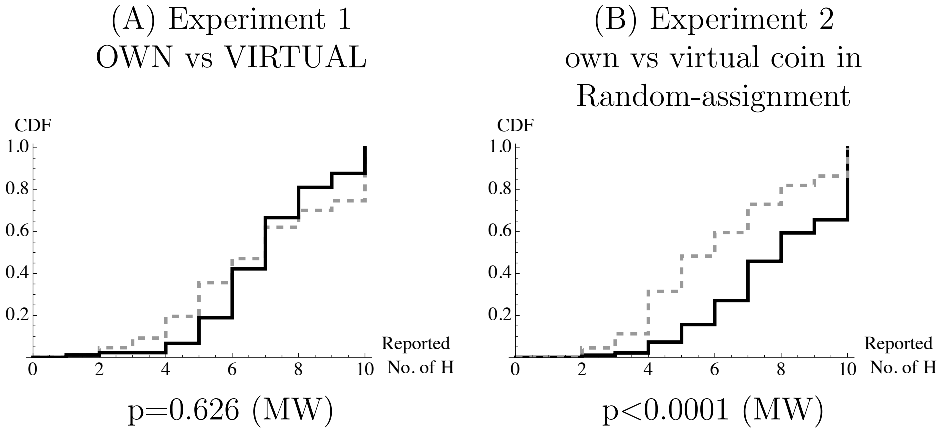


Figure 3: Distribution of the reported number of heads depending on the coin used, when the coin is assigned randomly. Own coin (Solid Black) vs Virtual coin (Dashed Gray). (A) Experiment 1: OWN vs VIRTUAL. (B) Experiment 2: Random-assignment procedure. MW: Mann-Whitney test, two-tailed. Holm-adjusted p-values are reported (two comparisons).

procedure in Experiment 2.[20]

In Panel (A), the reported number of heads does not differ significantly between the OWN and VIRTUAL treatments of Experiment 1. In Panel (B), by contrast, participants randomly assigned to use their own coin report significantly more heads than those randomly assigned to use the virtual coin. Thus, the observability effect documented in the literature (Kajackaite and Gneezy, 2017; Fries et al., 2021) is present in Experiment 2 but not in Experiment 1. Recall that participants in Experiment 1 were not aware that an alternative type of coin existed, whereas participants in Experiment 2 were aware of both types of coin and submitted a WTP to use their own coin. One possible explanation for the difference is salience: in Experiment 2, participants were explicitly aware of both coin options and submitted a WTP to use their own coin, making the difference in observability between the two coins more salient.[21]

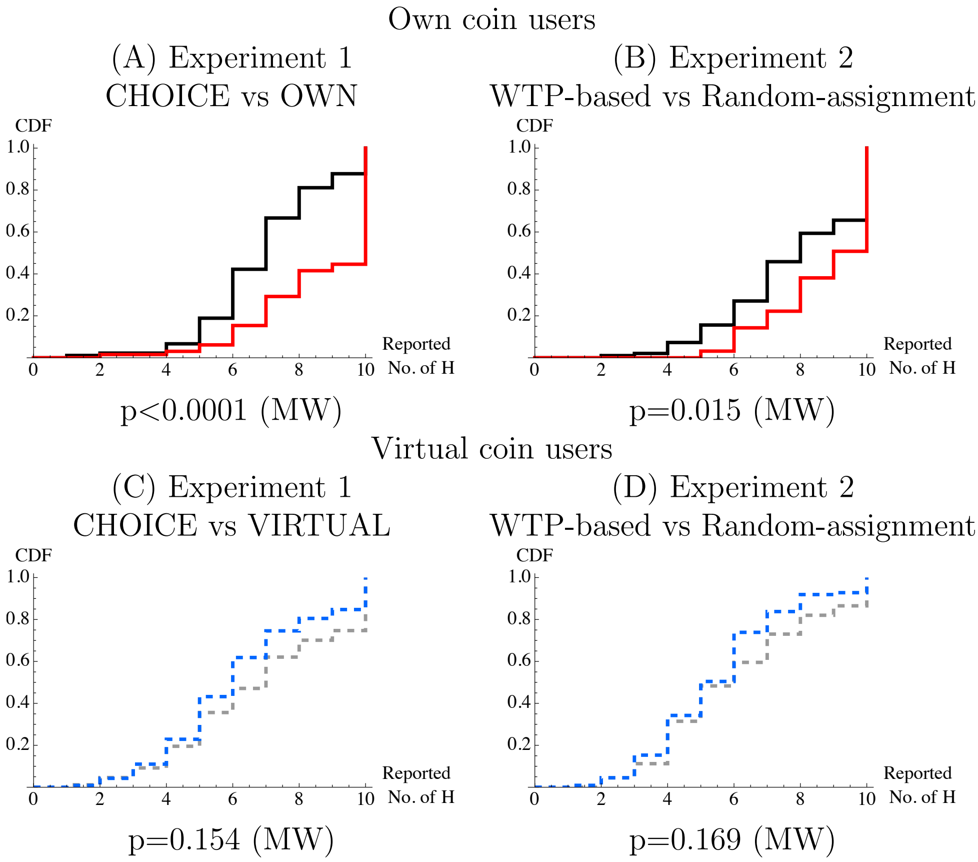


Figure 4: Distribution of the reported number of heads by treatments. (A) Own coin users in Experiment 1: CHOICE (Solid Red) vs OWN (Solid Black). (B) Own coin users in Experiment 2: WTP-based (Solid Red) vs Random-assignment (Solid Black). (C) Virtual coin users in Experiment 1: CHOICE (Dashed Blue) vs VIRTUAL (Dashed Gray). (D) Virtual coin users in Experiment 2: WTP-based (Dashed Blue) vs Random-assignment (Dashed Gray). MW: Mann-Whitney test, two-tailed. Holm-adjusted p-values are reported (four comparisons).

3.4 Selection and Licensing

Figure 4 compares CDFs of the reported number of heads by assignment procedure, separately for own-coin and virtual-coin users. Panels (A) and (B) compare reporting behavior of those who used their own coin depending on how they are assigned (based on WTP (Solid Red) or randomly (Solid Black)) in Experiments 1 and 2, respectively. Similarly, Panels (C) and (D) compare those who used the virtual coin depending on how they are assigned (based on WTP (Dashed Blue) or randomly (Dashed Gray)) in

[20] This analysis was not pre-registered.

[21] The absence of an observability effect in Experiment 1 may be due to a lack of power. Given our data from Experiment 1, the minimum detectable effect sizes (at the 5% significance level and 80% power) range from 0.77 to 0.89.

Experiments 1 and 2, respectively. We observe that those who used their own coin by paying a price to do so (in CHOICE or in the WTP-based procedure) report a significantly larger number of heads than those who were randomly assigned to use their own coin for free (in OWN or in Random-assignment) in both Experiments 1 and 2. However, among those who used the virtual coin, we do not observe significant differences in the reported number of heads depending on the assignment procedure in either experiment.

The significantly higher reported number of heads by own-coin users in the WTP-based procedure could be due to two mechanisms: "selection" (Erkal et al., 2011; Lazear et al., 2012; Grossman and van der Weele, 2017; Stüber, 2020) and "licensing".

Regarding selection, the higher the WTP submitted by participants in CHOICE or the WTP-based procedure, the higher the chance of using their own coin. Thus, participants with a higher WTP—and who may also be more inclined to misreport when unobserved—are more likely to select into using their own coin than those who are randomly assigned to use their own coin in OWN or in the Random-assignment procedure.

Regarding licensing, the fact that participants have paid a price to use their own coin may have justified their reporting a higher number of heads relative to those who are randomly assigned and do not pay anything to use their own coin.

Because Experiment 2 was designed explicitly to separate between these two mechanisms, we use data from that experiment alone to distinguish selection from licensing.

3.4.1 Selection

We first investigate the selection hypothesis which predicts that, among participants randomly assigned their own coin, those with a higher WTP to use their own coin report more heads. Table 3 reports OLS regressions of the reported number of heads on the submitted WTP and other covariates. The negative and significant coefficient on the "virtual coin" dummy (with own coin as the omitted category) confirms the observability effect documented above.

The estimated coefficients on WTP are positive but not significantly different from

Table 3: Selection: OLS regression results (Dependent variable is the reported number of heads)

	(1)	(2)	(3)	(4)
WTP	0.002	0.003**	0.002	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
virtual coin	-1.722***	-1.171**	-1.284***	-0.975*
	(0.322)	(0.547)	(0.321)	(0.509)
WTP× virtual coin		-0.003		-0.002
		(0.002)		(0.002)
Wuhan	0.792*	0.681	0.915*	0.843*
	(0.417)	(0.423)	(0.505)	(0.509)
Irvine	-0.177	-0.246	0.179	0.113
	(0.404)	(0.397)	(0.470)	(0.468)
No. Truthful			-0.184***	-0.181***
			(0.030)	(0.030)
No. WTP>0			-0.022	-0.023
			(0.032)	(0.032)
Constant	7.249***	7.043***	8.261***	8.192***
	(0.384)	(0.416)	(2.086)	(2.101)
Individual characteristics	No	No	Yes	Yes
Adjusted R^2	0.164	0.168	0.283	0.282
N	185	185	185	185
p: WTP+WTP×virtual coin=0		0.9881		0.6946

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Individual characteristics are gender, age, willingness to take risks, ethics and views on government monitoring; none is significant in any specification and coefficients are reported in Table D.3 in Online Appendix D.

zero in Columns (1) and (3). Thus, there is no significant relationship between WTP and reporting when this relationship is constrained to be the same for users of the two coins. Columns (2) and (4) allow the relationship to differ by coin type by including the interaction term WTP×virtual coin. Because own coin is the omitted category, the coefficient on WTP measures the relationship between WTP and reported heads among participants assigned their own coin. This coefficient is positive and significant in both specifications (2) and (4), providing evidence of selection among own-coin users. By contrast, among participants assigned the virtual coin, the sum of the coefficients on WTP and WTP×virtual coin is not significantly different from zero (see the bottom row of Table 3).

The absence of a selection effect among participants who use the virtual coin is consistent with our conceptual framework: WTP to use one's own coin should be higher

Table 4: Licensing: OLS regression results (Dependent variable is the reported number of heads)

	Own coin users		Own coin users who paid	
	WTP-based vs Random-assignment		CHOICE and WTP-based pooled	
	(1)	(2)	(3)	(4)
WTP	0.003***	0.002**	0.003*	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
WTP-based	0.352	0.373		
	(0.281)	(0.252)		
P			0.001	0.001
			(0.001)	(0.002)
Wuhan	0.443	0.988**	0.545**	0.774***
	(0.330)	(0.385)	(0.273)	(0.282)
Irvine	-0.831**	-0.083	-0.756*	-0.608
	(0.364)	(0.372)	(0.391)	(0.410)
No. Truthful		-0.163***		-0.103***
		(0.027)		(0.029)
No. WTP>0		0.027		0.022
		(0.034)		(0.044)
Constant	7.310***	7.637***	7.563***	10.301***
	(0.393)	(1.701)	(0.391)	(1.551)
Individual characteristics	No	Yes	No	Yes
Adjusted R^2	0.190	0.369	0.180	0.291
N	159	159	128	126

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Individual characteristics are gender, age, willingness to take risks, ethics and views on government monitoring. Female dummy is significant and negative in Column (2) but not in Column (4). Others are not significant in any specification. Coefficients are reported in Table D.4 in Online Appendix D.

among those who would benefit more from doing so rather than using the virtual coin.

3.4.2 Licensing

We next turn to the licensing hypothesis according to which, among those who used their own coin, the reported number of heads is higher for those who paid (i.e., in the WTP-based procedure) than those who did not (i.e., in the Random-assignment procedure), controlling for each participant's WTP to use their own coin.

Column (1) of Table 4 shows the results of regressing the reported number of heads on each participant's WTP to use their own coin, the dummy variable (WTP-based) that takes the value of 1 if the participant is in the WTP-based procedure and 0 if the participant is in the Random-assignment procedure, and location dummies. Column (2) further controls for individual characteristics and their beliefs about others' behavior.

We continue to observe a clear selection effect. The estimated coefficient on WTP is positive and significant in both specifications. However, we fail to observe a licensing effect. The estimated coefficient on the WTP-based dummy is positive but not statistically different from zero. We are, however, cautious about ruling the licensing effect out. Our design can rule out only effects of approximately 0.71–0.84 heads or larger, while the point estimates in Columns (1) and (2) are 0.35 and 0.37 heads.

We also conducted an additional analysis regarding licensing.[22] Conditional on the submitted WTP and on using one's own coin, the price paid, P, is randomly determined ($10 \leq P \leq WTP$). We exploit this variation to test whether participants who pay a higher price subsequently report more heads. To increase the sample size, we pool own-coin users from the WTP-based procedure in Experiment 2 and the CHOICE treatment in Experiment 1. Columns (3) and (4) of Table 4 report regressions of the number of heads on WTP, the price paid (P), and other controls. The estimated coefficient on P is not significantly different from zero in either specification. Thus, this additional analysis also provides no evidence in support of the licensing hypothesis.

4 Discussion and Conclusion

We have reported on an experiment that measures participants' willingness to pay for unobservability—that is, the ability to conceal misreporting in a coin-flipping task. We find that participants place substantial value on being unobservable: mean WTP is about 40% of the expected gain from maximum misreporting, and more than 90% of participants submit a positive WTP. This finding replicates almost exactly in Experiment 2, conducted three years later using the same subject pools.

It is important to be precise about what this WTP measures. Our conceptual framework distinguishes three forces. Monetary incentives encourage over-reporting; self-image concerns discourage it regardless of which coin is used; and privacy concerns reduce the cost of over-reporting when the realized outcome cannot be verified. Thus, in our experiment, WTP for unobservability is driven by privacy concerns. Social-image

[22] This additional analysis was not pre-registered.

concerns, however, work in the opposite direction. Because the experimenter observes each participant's submitted WTP, a positive WTP may itself be interpreted as signaling an intention to misreport. Participants who care about how they are perceived by the experimenter may therefore submit a lower WTP. Consequently, the WTP elicited in our experiment provides a lower bound on the value of unobservability.

Experiment 2 allows us to distinguish between two mechanisms that could explain why participants who pay for unobservability misreport more: selection and licensing. We find clear evidence for selection. Among participants randomly assigned their own coin for free, WTP is positively related to the number of heads reported, even though payment for the coin cannot itself affect their behavior. By contrast, WTP is unrelated to the number of heads reported among participants randomly assigned the virtual coin. This pattern is consistent with our conceptual framework: participants who value unobservability more are also those who report more heads when their outcomes cannot be verified. We do not find statistically significant evidence for licensing. Conditional on WTP, participants who pay to use their own coin do not report significantly more heads than participants who use their own coin for free. This conclusion should nevertheless be interpreted cautiously, since our design can rule out only licensing effects larger than about 0.8 reported heads.

A third result was not anticipated. The well-known observability effect in the literature appears in our data only when participants know that an unobserved alternative exists. Among participants simply handed one coin and never told of the other, reports are statistically indistinguishable regardless of whether or not the experimenter can verify them; among participants who had been asked to price unobservability and were then assigned a coin at random, those with their own coin report about one more head on average. The two comparisons differ in salience but also in the experiment in which they were run, so we regard this interpretation as suggestive.

While there are differences in WTP across the three locations, they are not stable. While WTP in Osaka was higher than in Wuhan in Experiment 1, the order was reversed, i.e., WTP in Wuhan was higher than in Osaka, in Experiment 2 conducted three years later. The three subject pools also differ in observable characteristics. Thus, we draw no cultural inference from these differences.

There is substantial variation in WTP in each location. WTP ranged from the minimum to the maximum possible value. To better understand the source of this variation, we examined a variety of individual characteristics. Among these variables, we found that participants' self-reported willingness to take risks is significantly and positively related to their WTP to use their own coin. By contrast, age, gender, and participants' views of ethical behavior are *not* related to WTP. Participants' views of the government's right to monitor people's activities were found to be positively related to WTP when we pool the data from two experiments. Interestingly, participants who are more accepting of the government's right to monitor economic activity are also willing to pay more to avoid being observed themselves.

We also observe that WTP is higher among those who believe that others are less likely to truthfully report the outcomes of their coin flips or those who believe others are more likely to report a positive WTP, suggesting that social norms may influence participants' social-image concerns (Gächter and Schulz, 2016; Serdarevic, 2021).

In future research, it would be of interest to consider other tasks with economic consequences where people face weaker or no material incentives to act dishonestly in a task in order to investigate whether such diminished motives matter for the WTP for unobservability. Similarly, it would be of interest to consider situations where dishonesty benefits others. While we suspect that there would be a reduction in the WTP in such settings, it could still be the case that individuals have a positive WTP for unobservability in their economic transactions that goes beyond the desire to avoid detection of cheating behavior.

Furthermore, the observed high values placed on unobservability in our experiment suggest a need for finding the right balance between fraud prevention and ensuring privacy in our digital economies. Effective fraud prevention, when balanced with privacy concerns, can enhance the functioning of the digital economy (Romanosky, 2016). Getting the right balance, however, can be a complex task because there are not only fraud prevention and privacy trade-offs but also trade-offs between different aspects of privacy itself (Pozen, 2016). Future studies might explore these varied trade-offs with the aim of discovering an optimal balance.

References

Abeler, J., D. Nosenzo, and C. Raymond (2019): "Preferences for Truth-Telling," *Econometrica*, 87, 1115–1153.

Acquisti, A., L. Brandimarte, and G. Loewenstein (2015): "Privacy and human behavior in the age of information," *Science*, 347, 509–514.

Acquisti, A., C. Taylor, and L. Wagman (2016): "The economics of privacy," *Journal of Economic Literature*, 54, 442–492.

Ahnert, T., P. Hoffmann, and C. Monnet (2022): "The digital economy, privacy, and CBDC," Working paper, ECB.

Becker, G. M., M. H. DeGroot, and J. Marschak (1964): "Measuring utility by a single-response sequential method," *Behavioral Science*, 9, 226–232.

Benndorf, V. and H.-T. Normann (2018): "The willingness to sell personal data," *The Scandinavian Journal of Economics*, 120, 1260–1278.

Beresford, A. R., D. Kübler, and S. Preibusch (2012): "Unwillingness to pay for privacy: A field experiment," *Economics Letters*, 117, 25–27.

Brebner, S. and J. Sonnemans (2018): "Does the elicitation method impact the WTA/WTP disparity?" *Journal of Behavioral and Experimental Economics*, 73, 40–45.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 90, 414–427.

Cason, T. N. and C. R. Plott (2014): "Misconceptions and game form recognition: Challenges to theories of revealed preference and framing," *Journal of Political Economy*, 122, 1235–1270.

Chen, D. L., M. Schonger, and C. Wicken (2016): "oTree – An open-source platform for laboratory, online and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Cialdini, R. B. and N. J. Goldstein (2004): "Social influence: Compliance and conformity," *Annual Review of Psychology*, 55, 591–621.

Cohn, A., E. Fehr, and M. A. Maréchal (2014): "Business culture and dishonesty in the banking industry," *Nature*, 516, 86–89.

Erkal, N., L. Gangadharan, and N. Nikiforakis (2011): "Relative earnings and giving in a real-effort experiment," *American Economic Review*, 101, 3330–3348.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018): "Global evidence on economic preferences," *Quarterly Journal of Economics*, 133, 1645–1692.

Fischbacher, U. and F. Föllmi-Heusi (2013): "Lies in disguise – An experimental study on cheating," *Journal of the European Economic Association*, 11, 525–547.

Fries, T., U. Gneezy, A. Kajackaite, and D. Parra (2021): "Observability and lying," *Journal of Economic Behavior and Organization*, 189, 132–149.

Gächter, S. and J. F. Schulz (2016): "Intrinsic honesty and the prevalence of rule violations across societies," *Nature*, 531, 496–499.

García-Gallego, A., N. Georgantzis, T. Jaber-López, and G. Michailidou (2020): "Audience effects and other-regarding preferences against corruption: Experimental evidence," *Journal of Economic Behavior & Organization*, 180, 159–173.

Gneezy, U., A. Kajackaite, and J. Sobel (2018): "Lying aversion and the size of the lie," *American Economic Review*, 108, 419–453.

Grossman, Z. (2015): "Self-signaling and social-signaling in giving," *Journal of Economic Behavior & Organization*, 117, 26–39.

Grossman, Z. and J. J. van der Weele (2017): "Self-image and willful ignorance in social decisions," *Journal of the European Economic Association*, 15, 173–217.

Haerpfer, C., R. Inglehart, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin, and B. Puranen, eds. (2022): *World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0.*, Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat.

Herskind, L., P. Katsikouli, and N. Dragoni (2020): "Privacy and cryptocurrencies—A systematic literature review," *IEEE Access*, 8, 54044–54059.

Jentzsch, N., S. Preibusch, and A. Harasser (2012): "Study on monetizing privacy: An economic model for pricing personal information," Technical report, European Network and Information Security Agency.

Kajackaite, A. and U. Gneezy (2017): "Incentives and cheating," *Games and Economic Behavior*, 102, 433–444.

Langer, E. J. (1975): "The illusion of control," *Journal of Personality and Social Psychology*, 32, 311–328.

Lazear, E. P., U. Malmendier, and R. A. Weber (2012): "Sorting in experiments with application to social preferences," *American Economic Journal: Applied Economics*, 4, 136–163.

MacKinnon, J. G. and M. D. Webb (2017): "Wild Bootstrap Inference for Wildly Different Cluster Sizes," *Journal of Applied Econometrics*, 32, 233–254.

Oshana, M. (2016): *Personal autonomy in society*, London, Routledge.

Pozen, D. E. (2016): "Privacy-privacy tradeoffs," *The University of Chicago Law Review*, 83, 221–247.

Preibusch, S., D. Kübler, and A. R. Beresford (2013): "Price versus privacy: An experiment into the competitive advantage of collecting less personal information," *Electronic Commerce Research*, 13, 432–455.

Prince, J. T. and S. Wallsten (2022): "How much is privacy worth around the world and across platforms?" *Journal of Economics & Management Strategy*, 31, 785–1019.

Roessler, B. and D. Mokrosinska (2013): "Privacy and social interaction," *Philosophy & Social Criticism*, 39, 771–791.

Romanosky, S. (2016): "Examining the costs and causes of cyber incidents," *Journal of Cybersecurity*, 2, 121–135.

Savage, S. J. and D. M. Waldman (2015): "Privacy tradeoffs in smartphone applications," *Economics Letters*, 137, 171–175.

Serdarevic, N. (2021): "Licence to lie and the social (In)appropriateness of lying," *Economics Letters*, 199, 109716.

Skatova, A., R. McDonald, S. Ma, and C. Maple (2023): "Unpacking privacy: Valuation of personal data protection," *PLoS One*, 18, e0284581.

Stüber, R. (2020): "The benefit of the doubt: Willful ignorance and altruistic punishment," *Experimental Economics*, 23, 848–872.

Tsai, J., S. Egelman, L. Cranor, and A. Acquisti (2011): "The effect of online privacy information on purchasing behavior: An experimental study," *Information Systems Research*, 22, 254–268.

Van Aaken, D., A. Ostermaier, and A. Picot (2014): "Privacy and Freedom: An Economic (Re-) Evaluation of Privacy," *Kyklos*, 67, 133–155.

Wang, Z. (2023): "Money Laundering and the Privacy Design of Central Bank Digital Currency," *Review of Economic Dynamics*, 51, 604–632.

Online Appendix
Paying to Be Unobservable: Evidence from a Lying Experiment

Te Bao* John Duffy† Nobuyuki Hanaki‡

A Instruction and Quiz (English version)

A.1 Instruction common to all

Welcome

- Welcome to the study. You are guaranteed $7 for showing up and completing this study.
- These instructions explain how you can earn *additional* earnings beyond the guaranteed $7 show-up payment from the decisions that you make.
- Additional earnings will be expressed in points. At the end of the experiment, any points you earned will be converted into US Dollar at the rate of 100 points = $ 1.00.
- Please silence any mobile devices and refrain from any distractions for the duration of this study. If you have any questions, please contact the experimenter.
- Today's study starts with the main decision task followed by a questionnaire. Your earning during the experiment will be paid in private.

*Nanyang Technological University

†University of California, Irvine, and Institute of Social and Economic Research, The University of Osaka

‡Corresponding author. Institute of Social and Economic Research, The University of Osaka, and University of Limassol

The main task

- The main task consists of flipping a two-sided coin 10 times. For each of the 10 flips, you record whether the coin landed heads or tails.
- After the completion of 10 flips, **you will receive 100 points for each head that you report and 0 points for each tail you report.**
- Note that your payment is based only on **the number of heads and tails** (out of 10 flips) that your report.

A.2 Instruction specific to CHOICE

The main task

- In completing this task, you have the option to use either
  - 1) a virtual online coin that we provide to you or
  - 2) to use your own coin.
- If you use your own coin, then it is not possible for us to check how many times your own coin actually landed heads.
- However, if you use our virtual coin, then we will be able to know the realization of the coin flips.
- **Still, regardless of whether you use your own coin or our virtual coin, we will rely on your *own report* of the number of heads and tails to determine your total points.**

The main task

- To use the virtual online coin costs you nothing.
- However, to use your own coin you have to first indicate your willingness to pay (WTP) in points for this option and in doing so, earn the right to use your own coin as explained in the next slide.

Willingness To Pay

- Specifically, we ask how much you are willing to pay (in points) to use your own coin.
- The range of values you can state is
  [0, 10, 20,..., 100, ... 200, ..., 300,... 400 ,..., 500],
  that is, the range is from 0 to 500 points in increments of 10 points.
- After you state your WTP, the computer program will draw a random price between 10 and 500 points, inclusive in increments of 10 points. Note that this is the same range for your stated WTP, except that 0 points is not included. All prices between 10 and 500 points, in increments of 10 points, are equally likely to be chosen.

Which coin to use?

- If your WTP is greater than or equal to the randomly drawn price, then you earn the right to use your own coin, but you have to pay the randomly drawn price in points.
- However, if the random price is greater than your stated WTP, then you do not earn the right to use your own coin and must use our virtual coin instead.
- Thus, the higher is your WTP, the greater is the chance that you will earn the right to use your own coin.
- At the extremes,
  - if your WTP is 0, then you will never earn the right to use your own coin,
  - if your WTP is 500, then you will always earn the right to use your own coin, but will have to pay the randomly drawn price out of your earnings.
- Of course, you are free to choose any WTP in between these two extremes.

two examples

- **Example 1**:
  - Imagine you have chosen your WTP to be 200 points. Then, the computer randomly draws a price of 100 points
  - In this case, you will use your own coin, but 100 points will be subtracted from your point earnings as the price for using your own coin.
- **Example 2**:
  - Imagine you have chosen your WTP to be 200 points. Then, the computer randomly draws a price of 300 points.
  - In this case, you must use our virtual coin, but no points will be subtracted from your point earnings.

How to determine your WTP?

- In deciding your WTP, please first ask yourself the following question:
  - At a price of 0 points, would you like to use your own coin instead of our virtual coin?
- If your answer is no, then your WTP is 0 points.
- If your answer is yes, then, go to the next question:
  - At the price of 10 points, would you like to use your own coin instead of our virtual coin?
- If your answer is no, then your WTP is 10 points.
- If your answer is yes, then, go to the next question:
  - At the price of 20 points, would you like to use your own coin instead of our virtual coin?
- **Continue this process until you find the price at which your answer switches from yes to no. This should be your WTP.**

Then, participants answered the following quiz.

A.3 Quiz for CHOICE

True or False

1 I will flip a coin 10 times and report the number of times the flip resulted in tails.

2 I can use the virtual coin without any cost.

3 If I submit some willingness to pay (WTP) amount, then I can always use my own coin to do the coin flips.

4 My payoff depends only on my own report of the outcome of the coin flips.

Choose the correct option

5-1 Consider the following scenario. You submit a WTP of 250 points and the random price drawn is 280 points. In this case:

a I use the virtual coin.

b I use my own coin.

5-2 If you report 5 heads, your payoff is:

a 250 points,

b 500 points,

c 780 points,

d 220 points

6-1 Consider the following scenario. You submit a WTP of 250 points and the random price drawn is 200 points. In this case:

a I use the virtual coin.

b I use my own coin.

6-2 If you report 5 heads, your payoff is:

a 250 points,

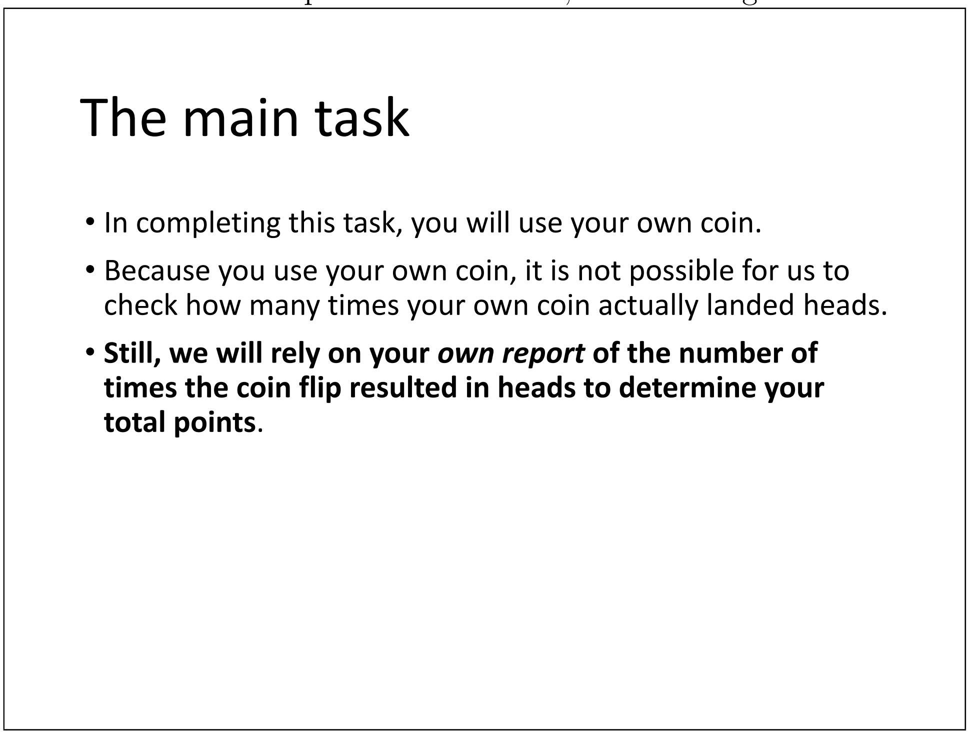
b 500 points,

c 700 points,

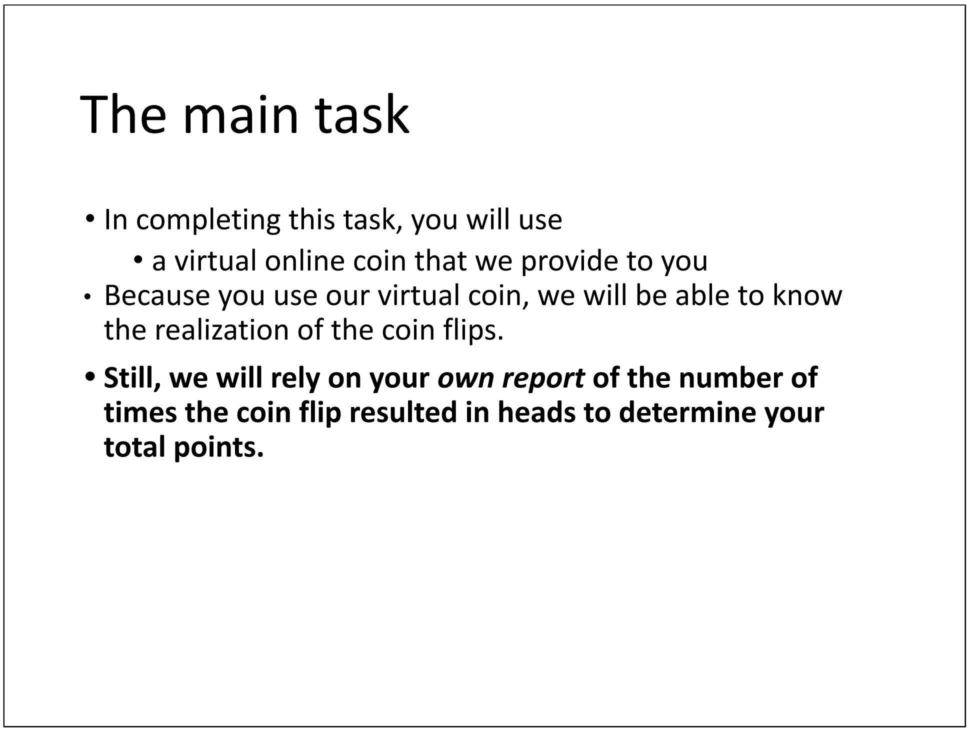
d 300 points

A.4 Instruction specific to OWN and VIRTUAL

Instead of the first slide specific to CHOICE, the following slide was used in OWN.



For VIRTUAL, the following slide was used.



No further instruction is given to participants. Participants answered the following quiz.

A.5 Quiz for OWN and VIRTUAL

True or False

1 I will flip a coin 10 times and report the number of times the flip resulted in tails.

2 My payoff depends only on my own report of the outcome of the coin flips.

A.6 Instruction specific to Experiment 2

After the final slide of the instruction to CHOICE, the following slide was added.

Determining the coin you use

- However, there is a chance that the coin you are going to use is determined randomly regardless of your choice of WTP and the price chosen by the computer.
- Namely, after you have submitted your WTP
  - with 50% probability the coin you use is determined randomly (random-assignment procedure)
    - Either your own coin or the virtual coin is selected with equal probability
    - In this case, you don't have to pay anything to use your own coin.
  - with 50% probability the coin you use is determined based on your WTP and the price chosen by computer as explained before (WTP-price procedure).
- You will be informed of which coin you will use, and how it is determined.

Then, participants answered the following quiz.

A.7 Quiz for Experiment 2

True or False

1 I will flip a coin 10 times and report the number of times the flip resulted in tails.

2 I can use the virtual coin without any cost.

3 If I submit some willingness to pay (WTP) amount, then I can always use my own coin to do the coin flips.

4 My payoff depends only on my own report of the outcome of the coin flips.

Choose the correct option

5-1 Consider the following scenario. Suppose the WTP-price procedure is randomly selected. You submit a WTP of 250 points and the random price drawn is 280 points. In this case:

a I use the virtual coin.

b I use my own coin.

5-2 If you report 5 heads, your payoff is:

a 250 points,

b 500 points,

c 780 points,

d 220 points

6-1 Consider the following scenario. Suppose the WTP-price procedure is randomly selected. You submit a WTP of 250 points and the random price drawn is 200 points. In this case:

a I use the virtual coin.

b I use my own coin.

6-2 If you report 5 heads, your payoff is:

a 250 points,

b 500 points,

c 700 points,

d 300 points

7 Why should you report your true WTP?

a Because you know for sure that the WTP-price procedure will be used.

b Because you do not know whether your WTP will matter, and if the WTP-price procedure is selected, your WTP determines which coin you use.

c Because reporting a higher WTP always increases your payoff.

d Because reporting zero WTP guarantees that you will use the virtual coin.

B Descriptive statistics

Table B.1: Mean (standard deviation) of participants' characteristics in each treatment and location. Original Experiment

Treatment	CHOICE				OWN				VIRTUAL			
	Wuhan	Irvine	Osaka	p-value[a]	Wuhan	Irvine	Osaka	p-value[a]	Wuhan	Irvine	Osaka	p-value[a]
No. obs.	60	60	63		30	30	30		30	30	27	
Proportion female	0.55 (0.50)	0.52 (0.50)	0.38 (0.49)	0.2275	0.73 (0.45)	0.70 (0.47)	0.33 (0.48)	0.0125	0.50 (0.51)	0.60 (0.50)	0.30 (0.47)	0.1354
Proportion others	0.07 (0.25)	0.02 (0.13)	0.02 (0.13)	0.2005	0.03 (0.18)	0.03 (0.18)	0.00 (0.00)	0.9676	0.07 (0.25)	0.00 (0.00)	0.00 (0.00)	0.8785
Age	21.12 (2.13)	21.58 (3.46)	22.60 (3.82)	0.0204	20.93 (1.91)	20.13 (2.49)	22.47 (2.86)	0.0004	21.4 (2.01)	20.6 (2.70)	22.63 (3.21)	0.0063
Risk-taking	4.13 (2.62)	5.58 (2.20)	4.40 (2.32)	0.0034	4.07 (2.45)	5.87 (2.01)	4.33 (2.43)	0.0037	3.63 (2.39)	5.33 (2.44)	4.33 (2.22)	0.0262
Ethics[b,d]	6.38 (1.89)	7.79 (1.31)	7.45 (1.35)	0.0001	6.29 (1.78)	7.30 (1.54)	7.99 (1.18)	0.0007	7.07 (1.86)	7.16 (1.66)	7.85 (1.30)	0.1239
Claiming benefit[d]	4.90 (2.89)	2.85 (1.94)	3.29 (2.51)	0.0002	5.17 (2.67)	3.73 (1.74)	3.31 (2.32)	0.0112	3.97 (2.57)	3.77 (2.14)	2.67 (2.17)	0.0240
Cheating tax[d]	3.36 (2.69)	2.73 (2.02)	2.57 (2.18)	0.3227	3.53 (2.50)	3.03 (1.97)	1.62 (0.82)	0.0023	2.57 (2.30)	3.33 (2.38)	2.11 (1.83)	0.0283
Telling truth[d]	5.40 (2.21)	6.93 (1.98)	6.21 (1.98)	0.0007	5.57 (1.77)	6.67 (1.97)	6.90 (2.11)	0.0080	5.73 (2.57)	6.57 (2.06)	6.33 (1.84)	0.2814
Government's right[c]	2.65 (0.67)	1.98 (0.53)	2.25 (0.71)	0.0001	2.78 (0.62)	2.22 (0.35)	2.33 (0.66)	0.0006	2.67 (0.69)	1.98 (0.63)	2.25 (0.70)	0.0006
Video surveillance	1.78 (0.85)	2.53 (0.82)	2.17 (0.87)	0.0001	1.70 (0.65)	2.00 (0.59)	2.20 (0.81)	0.0271	1.90 (0.92)	2.20 (0.96)	2.19 (1.00)	0.3851
Internet	2.68 (0.85)	3.44 (0.60)	3.10 (0.82)	0.0001	2.57 (0.86)	3.13 (0.68)	2.90 (0.76)	0.0195	2.60 (0.86)	3.37 (0.72)	3.15 (0.82)	0.0019
Bank account	2.58 (0.94)	3.14 (0.92)	2.98 (0.85)	0.0041	2.40 (0.93)	3.20 (0.61)	2.90 (0.88)	0.0019	2.50 (0.94)	3.50 (0.82)	2.93 (0.96)	0.0002
No. reporting truthfully	8.02 (6.05)	9.23 (4.95)	8.25 (5.04)	0.3640	6.23 (5.20)	9.83 (3.57)	9.50 (6.10)	0.0088	5.90 (6.32)	10.90 (5.74)	10.11 (6.44)	0.0046
No. WTP>0	13.5 (5.89)	12.78 (4.92)	15.35 (5.14)	0.0095								
Payoff	714.33 (238.93)	619.00 (214.42)	626.19 (205.36)	0.0355	796.67 (158.62)	636.67 (140.16)	640.00 (201.03)	0.0005	776.67 (256.88)	636.67 (245.63)	607.41 (218.26)	0.0151

[a]: P-values are for comparison across countries (Kruskal-Wallis test). These p-values are descriptive checks on the composition of the samples rather than tests of the study's hypotheses, and are not adjusted for multiple comparisons.
[b]: Ethics = ((11 - claiming benefit) + (11 - cheating tax) + telling truth)/3.
[c]: Government's right = ((5-video surveillance) + (5-internet) + (5-bank account))/3.
[d]: Due to a technical problem, some answers were not recorded; thus n=58 in Wuhan (CHOICE), n=59 in Irvine (CHOICE), n=29 in Osaka (OWN)

Table B.2: Mean (standard deviation) of participants' characteristics in each condition and location. Additional Experiment

Treatment	WTP-based				Random-assignment: Own coin				Random-assignment: Virtual coin			
	Wuhan	Irvine	Osaka	p-value[a]	Wuhan	Irvine	Osaka	p-value[a]	Wuhan	Irvine	Osaka	p-value[a]
No. obs.	60	57	57		29	34	33		31	29	29	
Proportion female	0.73 (0.45)	0.65 (0.48)	0.35 (0.48)	0.0009	0.72 (0.45)	0.62 (0.49)	0.39 (0.50)	0.0703	0.58 (0.50)	0.62 (0.49)	0.28 (0.45)	0.0458
Proportion other	0.03 (0.18)	0.05 (0.23)	0.02 (0.13)	0.9488	0.03 (0.19)	0.00 (0.00)	0.00 (0.00)	0.9649	0.00 (0.00)	0.00 (0.00)	0.03 (0.19)	0.9661
Age	21.12 (2.08)	20.35 (2.18)	22.14 (2.64)	0.0001	20.69 (1.81)	20.26 (2.00)	22.03 (3.07)	0.0343	21.19 (1.85)	20.28 (1.53)	21.90 (1.72)	0.0027
Risk-taking	5.26 (2.39)	5.58 (1.84)	3.82 (2.25)	0.0043	4.31 (2.12)	6.03 (2.07)	4.82 (2.14)	0.0079	4.39 (2.79)	5.69 (2.19)	4.41 (2.40)	0.0441
Ethics[b]	7.12 (1.48)	7.56 (1.32)	7.91 (1.02)	0.0141	6.80 (1.45)	7.35 (1.21)	7.76 (1.17)	0.0372	6.95 (1.64)	7.70 (1.11)	7.87 (1.28)	0.0733
Claiming benefit	3.98 (2.30)	3.49 (2.05)	2.40 (1.77)	0.0002	4.58 (2.46)	3.59 (1.91)	2.70 (1.99)	0.0049	4.32 (2.65)	2.93 (1.56)	2.79 (2.29)	0.0295
Cheating tax	2.47 (2.04)	3.14 (1.92)	1.84 (1.15)	0.0005	2.52 (1.92)	2.82 (1.82)	1.91 (1.44)	0.0768	2.65 (2.35)	2.76 (1.62)	1.79 (1.52)	0.0594
Telling truth	5.83 (1.86)	7.30 (1.82)	5.98 (2.40)	0.0002	5.52 (1.81)	6.47 (2.16)	5.88 (2.10)	0.1894	5.81 (2.24)	6.79 (1.84)	6.21 (2.34)	0.1815
Government's right[c]	2.84 (0.71)	1.81 (0.61)	2.18 (0.65)	0.0001	2.66 (0.69)	2.01 (0.65)	2.35 (0.55)	0.0004	2.81 (0.64)	1.83 (0.52)	2.41 (0.74)	0.0001
Video surveillance	1.72 (0.78)	2.65 (1.01)	2.21 (0.77)	0.0001	1.79 (0.77)	2.56 (0.96)	2.00 (0.79)	0.0048	1.81 (0.65)	2.48 (0.83)	1.97 (0.91)	0.0129
Internet	2.47 (0.91)	3.58 (0.63)	3.21 (0.77)	0.0001	2.62 (0.78)	3.29 (0.68)	3.15 (0.76)	0.0050	2.48 (0.96)	3.45 (0.57)	2.97 (0.94)	0.0008
Bank account	2.28 (0.96)	3.35 (0.69)	3.04 (0.80)	0.0001	2.62 (0.98)	3.12 (0.91)	2.79 (0.86)	0.1275	2.29 (0.94)	3.59 (0.63)	2.83 (1.00)	0.0001
No. reporting truthfully	7.8 (4.66)	10.47 (4.92)	7.67 (5.94)	0.0037	5.83 (4.36)	8.94 (5.34)	6.00 (5.68)	0.0205	9.90 (4.64)	10.00 (5.20)	9.24 (4.31)	0.8225
No. WTP>0	15.1 (4.59)	13.07 (5.15)	15.39 (5.30)	0.0131	16.28 (4.36)	14.03 (4.31)	15.18 (5.60)	0.0563	11.71 (5.29)	13.38 (4.79)	13.21 (6.33)	0.3468
Payoff	614.17 (199.49)	561.40 (183.57)	616.49 (213.66)	0.2645	879.31 (152.08)	708.82 (178.15)	754.55 (243.79)	0.0036	632.26 (252.17)	606.90 (213.69)	568.97 (233.15)	0.4987
WTP	223.67 (170.22)	152.63 (118.86)	223.68 (169.14)	0.0700	259.31 (173.27)	210.59 (148.32)	107.28 (83.38)	0.0005	158.39 (143.46)	170.69 (127.22)	161.03 (135.31)	0.8065

[a]: P-values are for comparison across countries (Kruskal-Wallis test). These p-values are descriptive checks on the composition of the samples rather than tests of the study's hypotheses, and are not adjusted for multiple comparisons.
[b]: Ethics = ((11 - claiming benefit) + (11 - cheating tax) + telling truth)/3.
[c]: Government's right = ((5-video surveillance) + (5-internet) + (5-bank account))/3.

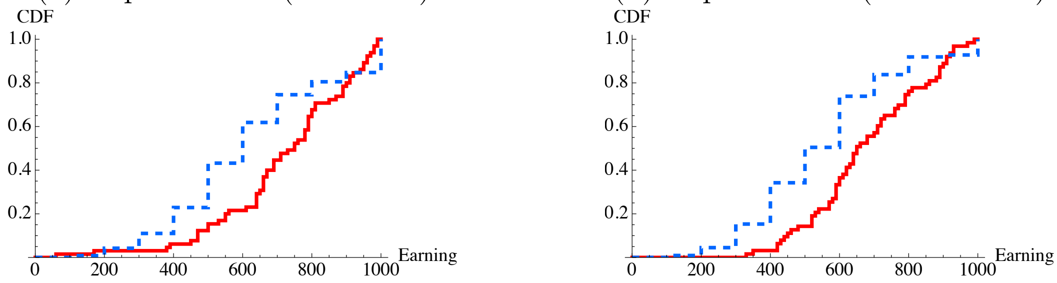
Table B.3: Mean (standard deviation) of participants' characteristics in each condition.

Treatment	Original Experiment				Additional Experiment				
	CHOICE	OWN	VIRTUAL	p-value[a]	WTP-based	Random own	Random virtual	p-value[b]	p-value[c]
No. obs.	183	90	87		174	96	89		
Proportion female	0.48 (0.50)	0.59 (0.49)	0.47 (0.50)	0.2852	0.58 (0.49)	0.57 (0.50)	0.49 (0.50)	0.4958	0.3478
Proportion other	0.03 (0.18)	0.02 (0.15)	0.02 (0.15)	0.9861	0.03 (0.18)	0.01 (0.10)	0.01 (0.11)	0.9275	0.9992
Age	21.78 (3.27)	21.18 (2.61)	21.51 (2.76)	0.2255	21.20 (2.41)	21.00 (2.48)	21.12 (1.81)	0.3751	0.1397
Risk-taking	4.70 (2.45)	4.76 (2.41)	4.44 (2.43)	0.5750	4.52 (2.24)	5.09 (2.21)	4.82 (2.53)	0.1790	0.4275
Ethics[d,f]	7.21 (1.64)	7.18 (1.66)	7.34 (1.65)	0.8001	7.52 (1.32)	7.33 (1.32)	7.49 (1.41)	0.3449	0.6608
Claiming benefit[f]	3.66 (2.61)	4.08 (2.38)	3.49 (2.35)	0.1386	3.30 (2.15)	3.58 (2.23)	3.37 (2.31)	0.5794	0.2132
Cheating tax[f]	2.88 (2.32)	2.74 (2.05)	2.69 (2.23)	0.9121	2.48 (1.82)	2.42 (1.76)	2.40 (1.91)	0.7566	0.7032
Telling truth[f]	6.18 (2.14)	6.37 (2.02)	6.21 (2.19)	0.8766	6.36 (2.13)	5.98 (2.06)	6.26 (2.17)	0.4106	0.8422
Government's right[e]	2.29 (0.70)	2.44 (0.61)	2.30 (0.73)	0.2200	2.29 (0.79)	2.32 (0.68)	2.36 (0.75)	0.7815	0.6264
Video surveillance	2.16 (0.89)	1.97 (0.71)	2.09 (0.96)	0.3551	2.18 (0.94)	2.14 (0.90)	2.08 (0.84)	0.7830	0.7275
Internet	3.07 (0.82)	2.87 (0.80)	3.03 (0.86)	0.1871	3.07 (0.91)	3.04 (0.78)	2.96 (0.93)	0.6235	0.4622
Bank account	2.90 (0.93)	2.83 (0.88)	2.98 (0.99)	0.5486	2.88 (0.94)	2.85 (0.93)	2.89 (1.02)	0.9283	0.9223
No. reporting truthfully	8.50 (5.36)	8.52 (5.27)	8.93 (6.49)	0.9264	8.63 (5.32)	6.99 (5.34)	9.72 (4.69)	0.0006	0.0172
No. WTP>0	13.90 (5.41)				14.53 (5.09)	15.10 (4.84)	12.74 (5.49)	0.0054	0.0125
Payoff	652.73 (222.82)	691.11 (182.77)	675.86 (250.10)	0.2775	597.64 (199.76)	776.04 (207.11)	603.37 (232.79)	0.0001	0.0001
WTP	188.09 (143.68)				200.40 (157.59)	189.79 (151.09)	163.26 (134.24)	0.2481	0.4044

a: P-values are for comparison across three treatments in Experiment 1 (Kruskal-Wallis test).
b: P-values are for comparison across three conditions in Experiment 2 (Kruskal-Wallis test).
c: P-values are for comparison across all the conditions (Kruskal-Wallis test).
These Kruskal-Wallis p-values are descriptive checks on the composition of the samples rather than tests of the study's hypotheses, and are not adjusted for multiple comparisons.
d: Ethics = ((11 - claiming benefit) + (11 - cheating tax) + telling truth)/3.
e: Government's right = ((5-video surveillance) + (5-internet) + (5-bank account))/3.
f: Due to a technical problem, some answers were not recorded; thus n=180 in CHOICE, n=89 in OWN.

C Payoffs

(a) Experiment 1 (CHOICE) (b) Experiment 2 (WTP-based)



	Own coin	Virtual coin	p-value
Mean	719.23	616.10	p=0.0018
(s.d.)	(195.68)	(229.07)	

	Own coin	Virtual coin	p-value
Mean	677.62	552.25	<0.0001
(s.d.)	(165.32)	(203.98)	

p-value based on Mann-Whitney test, two-tailed. Holm corrected (two comparisons).

Figure C.1: Cumulative Distributions of payoffs depending on the coin used. Own coin (Solid Red) and the virtual coin (Dashed Blue). (a) CHOICE treatment of Experiment 1. (b) WTP-based procedure in Experiment 2.

Figure C.1 shows the cumulative distribution of payoffs obtained by those who used their own coin (Solid Red) and by those who used the virtual coin (Dashed Blue) in the CHOICE treatment of Experiment 1 (panel a) and in the WTP-based procedure of Experiment 2 (panel b).

Those who used their own coin earned significantly more than those who used the virtual coin (see also the negative and significant coefficient of the dummy variable "Virtual coin" which takes the value 1 if a participant used the virtual coin, and 0 otherwise, in Table C.1). One could argue that those who purchased the right to use their own coin were "rational" in doing so, at least in terms of material payoff.

Table C.1: Payoffs

	Exp 1 (CHOICE)	Exp 2 (WTP-based)	Pooled
Wuhan	95.904**	-21.218	95.904**
	(39.636)	(40.485)	(39.665)
Irvine	4.229	-36.648	4.229
	(41.562)	(42.127)	(41.591)
Exp 2			21.568
			(266.563)
Exp 2 × Wuhan			-117.121**
			(56.656)
Exp 2 × Irvine			-40.876
			(59.176)
WTP	-0.219	-0.284*	-0.219
	(0.203)	(0.144)	(0.203)
Virtual coin	-156.172**	-120.740*	-156.172**
	(68.363)	(65.938)	(68.412)
WTP × Virtual coin	0.290	-0.149	0.290
	(0.276)	(0.323)	(0.277)
Exp 2 × WTP			-0.066
			(0.249)
Exp 2 × Virtual coin			35.432
			(94.982)
Exp 2 × WTP × Virtual coin			-0.439
			(0.425)
Female	33.901	1.375	33.901
	(32.577)	(30.984)	(32.601)
Other	9.485	65.682	9.485
	(69.689)	(125.869)	(69.739)
Age	1.705	-3.713	1.705
	(5.236)	(6.023)	(5.240)
Risk	1.976	3.985	1.976
	(7.795)	(7.047)	(7.801)
Ethics	1.313	-2.975	1.313
	(9.805)	(10.577)	(9.812)
Gov. right	-2.656	24.749	-2.656
	(22.729)	(19.575)	(22.745)
No. Truthful	-16.518***	-12.096***	-16.518***
	(2.964)	(3.258)	(2.967)
No. WTP>0	-3.183	0.624	-3.183
	(3.708)	(3.403)	(3.711)
Exp 2 × Female			-32.526
			(44.960)
Exp 2 × Other			56.198
			(143.816)
Exp 2 × Age			-5.418
			(7.980)
Exp 2 × Risk			2.008
			(10.509)
Exp 2 × Ethics			-4.288
			(14.421)
Exp 2 × Gov. right			27.406
			(29.999)
Exp 2 × No. Truthful			4.422
			(4.405)
Exp 2 × No. WTP>0			3.807
			(5.033)
Constant	859.010***	880.578***	859.010***
	(170.629)	(204.847)	(170.751)
Adjusted R^2	0.198	0.175	0.200
N	180	174	354
H_0: Wuhan = Irvine	0.0375	0.6737	
H_0: location × Exp 2 = 0			0.1088
H_0: characteristics, beliefs × Exp 2 = 0			0.8761

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
In the Pooled column, all coefficients are allowed to differ between the two experiments, so the Experiment 1 coefficients are reproduced exactly and the Experiment 2 effects are read off the interaction terms.

D Supporting regressions for Tables 1 to 4 in the main text

Table D.1: Willingness to pay (WTP) to use own coin: unrestricted specification underlying the joint tests reported in Table 1

Wuhan	-61.625**
	(25.097)
Irvine	-39.090
	(26.660)
Exp 2	-67.121
	(127.012)
Exp 2 × Wuhan	106.480***
	(33.610)
Exp 2 × Irvine	63.671*
	(33.679)
Female	-5.315
	(19.707)
Other	52.420
	(45.286)
Age	0.570
	(3.307)
Risk	16.611***
	(4.401)
Ethics	-3.539
	(5.968)
Gov. right	17.714
	(14.841)
No. Truthful	-5.563***
	(1.703)
No. WTP>0	6.472***
	(1.889)
Exp 2 ×Female	-20.493
	(25.536)
Exp 2 ×Other	-111.804
	(76.115)
Exp 2 ×Age	-1.028
	(4.857)
Exp 2 ×Risk	-10.226*
	(5.694)
Exp 2 ×Ethics	5.977
	(8.469)
Exp 2 ×Gov. right	-1.597
	(18.870)
Exp 2 ×No. Truthful	1.855
	(2.327)
Exp 2 ×No. WTP>0	2.606
	(2.405)
Constant	73.948
	(95.056)
Adjusted R^2	0.166
N	539

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Table D.1, all coefficients are allowed to differ between the two experiments; the joint tests in the final two rows of Table 1 are computed from this regression. Column (6) of Table 1 imposes equality of the coefficients on individual characteristics and beliefs — a restriction these tests do not reject — and its coefficients therefore differ from those reported here.

Table D.2: Reported number of heads and willingness to pay: supporting specifications for Table 2

	Exp 1 (CHOICE)	Exp 2 (WTP-based)	Pooled Unrestricted	
	(2)	(4)	(a)	(b)
Wuhan	0.981**	-0.271	1.014***	0.981**
	(0.381)	(0.404)	(0.375)	(0.381)
Irvine	-0.090	-0.561	-0.073	-0.090
	(0.403)	(0.394)	(0.380)	(0.403)
Exp 2			1.048	0.497
			(0.764)	(2.580)
Exp 2 × Wuhan			-1.059**	-1.253**
			(0.516)	(0.555)
Exp 2 × Irvine			-0.761	-0.471
			(0.520)	(0.563)
WTP	0.002	0.003***	0.005***	0.002
	(0.002)	(0.001)	(0.001)	(0.002)
Virtual coin	-1.519**	-1.035	-0.943	-1.519**
	(0.626)	(0.637)	(0.638)	(0.626)
WTP × Virtual coin	-0.001	-0.008**	-0.006**	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)
Exp 2 × WTP			-0.002	0.001
			(0.002)	(0.002)
Exp 2 × Virtual coin			-0.804	0.484
			(0.881)	(0.893)
Exp 2 × WTP × Virtual coin			-0.000	-0.006
			(0.004)	(0.004)
Female	0.250	-0.069		0.250
	(0.314)	(0.297)		(0.314)
Other	-0.214	0.272		-0.214
	(0.578)	(1.176)		(0.579)
Age	0.034	-0.022		0.034
	(0.047)	(0.058)		(0.047)
Risk	0.011	0.018		0.011
	(0.074)	(0.067)		(0.074)
Ethics	0.030	-0.084		0.030
	(0.098)	(0.104)		(0.098)
Gov. right	-0.097	0.231		-0.097
	(0.217)	(0.190)		(0.217)
No. Truthful	-0.176***	-0.109***		-0.176***
	(0.028)	(0.031)		(0.028)
No. WTP>0	-0.035	0.001		-0.035
	(0.035)	(0.033)		(0.035)
Exp 2 × Female				-0.319
				(0.432)
Exp 2 × Other				0.486
				(1.310)
Exp 2 × Age				-0.056
				(0.075)
Exp 2 × Risk				0.007
				(0.100)
Exp 2 × Ethics				-0.114
				(0.143)
Exp 2 × Gov. right				0.327
				(0.289)
Exp 2 × No. Truthful				0.067
				(0.042)
Exp 2 × No. WTP>0				0.036
				(0.048)
Constant	8.448***	8.945***	6.891***	8.448***
	(1.624)	(2.006)	(0.514)	(1.625)
Adjusted R^2	0.384	0.470	0.348	0.427
N	180	174	357	354

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Columns (2) and (4) of Table D.2 reproduce the corresponding columns of Table 2 with the individual characteristics shown. Columns (a) and (b) allow every coefficient to differ between the two experiments and are the basis for the joint tests reported in Table 2; they correspond to the pooled specifications in Columns (5) and (6) of Table 2, which impose the restrictions those tests do not reject.

Table D.3: Selection: Full specification of Columns (3) and (4) of Table 3

	(3)	(4)
WTP	0.002	0.002**
	(0.001)	(0.001)
virtual coin	-1.284***	-0.975*
	(0.321)	(0.509)
WTP× virtual coin		-0.002
		(0.002)
Wuhan	0.915*	0.843*
	(0.505)	(0.509)
Irvine	0.179	0.113
	(0.470)	(0.468)
Female	0.013	0.036
	(0.353)	(0.355)
Other	-0.396	-0.385
	(0.734)	(0.819)
Age	-0.009	-0.011
	(0.071)	(0.071)
Risk	-0.004	-0.000
	(0.073)	(0.073)
Ethics	0.089	0.088
	(0.117)	(0.116)
Gov. right	0.001	-0.007
	(0.247)	(0.245)
No. Truthful	-0.184***	-0.181***
	(0.030)	(0.030)
No. WTP>0	-0.022	-0.023
	(0.032)	(0.032)
Constant	8.261***	8.192***
	(2.086)	(2.101)
Adjusted R^2	0.283	0.282
N	185	185

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

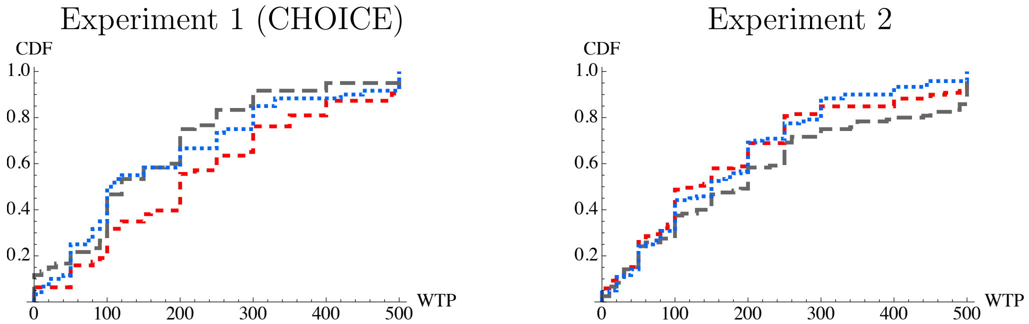
Table D.4: Licensing: Full specification for Columns (2) and (4) of Table 4

	Own coin users (2)	Own coin users who paid (4)
WTP	0.002**	0.002
	(0.001)	(0.001)
WTP-based	0.373	
	(0.252)	
P		0.001
		(0.002)
Wuhan	0.988**	0.774***
	(0.385)	(0.282)
Irvine	-0.083	-0.608
	(0.372)	(0.410)
Female	-0.635**	-0.385
	(0.305)	(0.296)
Other	0.092	-0.454
	(0.342)	(0.525)
Age	0.014	-0.047
	(0.058)	(0.043)
Risk	0.028	-0.005
	(0.055)	(0.064)
Ethics	0.033	-0.071
	(0.107)	(0.096)
Gov. right	-0.064	-0.176
	(0.202)	(0.194)
No. Truthful	-0.163***	-0.103***
	(0.027)	(0.029)
No. WTP>0	0.027	0.022
	(0.034)	(0.044)
Constant	7.637***	10.301***
	(1.701)	(1.551)
Adjusted R^2	0.369	0.291
N	159	126

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Location-by-location analyses

E.1 WTP to use own coin



	Wuhan	Irvine	Osaka
Mean	159.83	177.00	225.56
(s.d.)	(126.81)	(146.20)	(150.48)

	Wuhan	Irvine	Osaka
Mean	251.42	173.42	176.13
(s.d.)	(167.24)	(131.00)	(199.27)

Figure E.1: Cumulative distributions of willingness to pay to use own coin in Osaka (red dashed), Irvine (cyan dots), and Wuhan (long gray dash).

Figure E.1 shows the cumulative distribution of WTP to use own coin in Osaka (red dashed), Irvine (cyan dots), and Wuhan (long gray dash) in Experiments 1 (left) and 2 (right). As noted in the main text, WTP to use own coin is significantly lower in Wuhan than in Osaka in Experiment 1, while the opposite is the case in Experiment 2.

Location-by-location regression results reported in Table E.1 show some variations in correlations between individual characteristics and WTP. Among those characteristics that are significantly correlated with WTP in the pooled regression reported in Table 1, the willingness to take risks is significantly correlated with WTP in Wuhan and Irvine but not in Osaka, and the beliefs about others' reporting behavior are significantly correlated with WTP in Osaka and Wuhan, but not in Irvine.

In Columns (3), (6), and (9) of Table E.1, all coefficients are allowed to differ between the two experiments, so the Experiment 1 coefficients are reproduced exactly and the Experiment 2 effects are read off the interaction terms.

The final row of Table E.1 tests, location by location, whether the coefficients on individual characteristics and beliefs differ between the two experiments. Equality is not rejected in Osaka (p=0.41) or Wuhan (p=0.47) but is rejected in Irvine ($p<0.0001$), although the pooled test does not reject (p=0.48, Table 1). We have no explanation for this, and take it as a further reason to treat location-level results as descriptive.

Table E.1: Willingness to pay (WTP) to use own coin: Location-by-location OLS regression

	Osaka			Wuhan			Irvine		
	Exp 1	Exp 2	Pooled	Exp 1	Exp 2	Pooled	Exp 1	Exp 2	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Female	-1.134	-40.531	-1.134	10.655	-9.509	10.655	-34.892	-18.147	-34.892
	(35.576)	(26.600)	(34.698)	(30.209)	(31.950)	(29.287)	(33.164)	(23.903)	(32.191)
Other	24.047	72.334	24.047	9.413	-144.499***	9.413	175.434***	-15.815	175.434***
	(60.885)	(183.847)	(59.381)	(54.896)	(48.533)	(53.220)	(34.583)	(33.019)	(33.569)
Age	8.867*	8.772**	8.867**	-13.662**	-5.902	-13.662**	-8.421	-6.448	-8.421
	(4.440)	(4.329)	(4.330)	(6.310)	(7.936)	(6.118)	(6.834)	(5.015)	(6.633)
Risk	1.947	-0.773	1.947	11.466**	8.355	11.466**	37.513***	13.616***	37.513***
	(8.988)	(7.132)	(8.766)	(4.994)	(6.763)	(4.841)	(9.158)	(5.185)	(8.890)
Ethics	-12.742	12.343	-12.742	-2.767	12.726	-2.767	4.462	-13.153*	4.462
	(12.297)	(12.725)	(11.994)	(8.519)	(9.948)	(8.258)	(11.210)	(7.486)	(10.881)
Gov. right	11.867	34.923	11.867	16.047	-2.169	16.047	22.053	10.473	22.053
	(24.231)	(21.665)	(23.632)	(21.712)	(23.441)	(21.050)	(35.026)	(17.136)	(33.999)
No. Truthful	-11.482***	-5.447**	-11.482***	-6.609**	-7.458**	-6.609**	-0.182	-1.090	-0.182
	(3.351)	(2.443)	(3.268)	(2.800)	(3.206)	(2.714)	(3.047)	(2.386)	(2.958)
No. WTP>0	10.912***	6.927***	10.912***	9.267***	10.448***	9.267***	0.374	11.303***	0.374
	(3.573)	(2.261)	(3.485)	(2.140)	(3.419)	(2.075)	(3.306)	(2.058)	(3.209)
Exp 2			-251.417			-162.045			97.613
			(210.682)			(220.511)			(185.818)
Exp 2 × Female			-39.397			-20.165			16.745
			(43.929)			(43.683)			(40.297)
Exp 2 × Other			48.287			-153.912**			-191.249***
			(195.445)			(72.501)			(47.414)
Exp 2 × Age			-0.096			7.760			1.973
			(6.163)			(10.111)			(8.359)
Exp 2 × Risk			-2.720			-3.111			-23.898**
			(11.359)			(8.396)			(10.328)
Exp 2 × Ethics			25.086			15.493			-17.615
			(17.606)			(13.040)			(13.268)
Exp 2 × Gov. right			23.056			-18.216			-11.580
			(32.249)			(31.757)			(38.183)
Exp 2 × No. Truthful			6.036			-0.849			-0.908
			(4.099)			(4.236)			(3.821)
Exp 2 × No. WTP>0			-3.985			1.181			10.929***
			(4.170)			(4.042)			(3.828)
Constant	12.150	-239.267*	12.150	298.991**	136.946	298.991**	81.301	178.913*	81.301
	(168.478)	(130.189)	(164.317)	(144.650)	(167.756)	(140.233)	(154.941)	(107.611)	(150.397)
Adjusted R^2	0.166	0.106	0.143	0.342	0.143	0.196	0.232	0.237	0.234
N	63	119	182	58	120	178	59	120	179
p-value[a]			0.4071			0.4675			<0.0001

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
a: H_0: characteristics, beliefs × Exp 2 = 0

E.2 Reporting behavior: Observability Effect

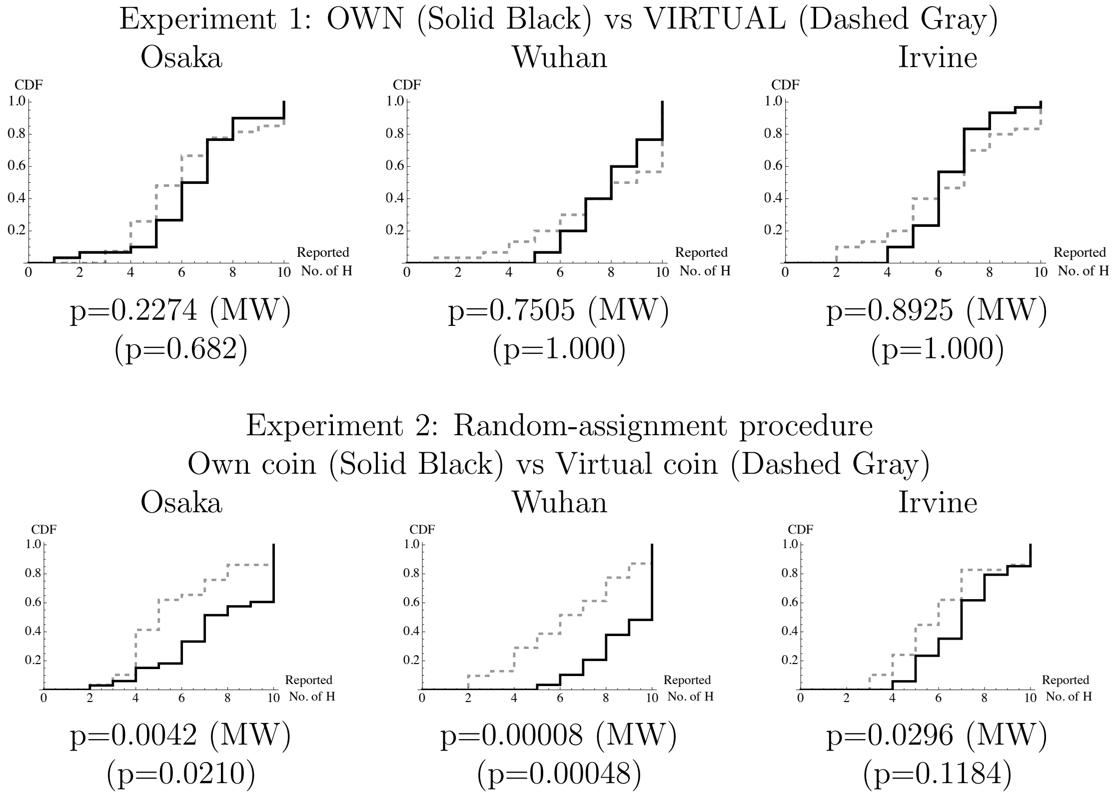


Figure E.2: CDFs of reported number of heads depending on the randomly assigned coin used: Own coin (Solid Black) and the virtual coin (Dashed Gray). Experiment 1 compares the OWN and VIRTUAL treatments. Experiment 2 compares own-coin and virtual-coin users in the Random-assignment procedure. p-values are from Mann-Whitney tests. p-values in parentheses are based on Holm correction (six comparisons).

Figure E.2 shows the CDFs of the reported number of heads among those who are randomly assigned their own coin (Solid Black) and the virtual coin (Dashed Gray) in Experiments 1 (top) and 2 (bottom) in Osaka, Wuhan, and Irvine. Regardless of the location, there is no detectable observability effect in Experiment 1, while it is clearly observed in Experiment 2 (except in Irvine). Thus, the salience of the availability of an unobservable option seems important regardless of the location.

E.3 Reporting behavior: Selection and Licensing

Figure E.3 shows the CDFs of the reported number of heads depending on the way the coin is assigned: WTP-based (colored) vs random assignment (grayscale). Among own coin users, WTP-based (Solid Red) and randomly assigned (Solid Black) are compared in each experiment. Similarly, among the virtual coin users, WTP-based (Dashed Blue) and randomly assigned (Dashed Gray) are compared in each experiment.

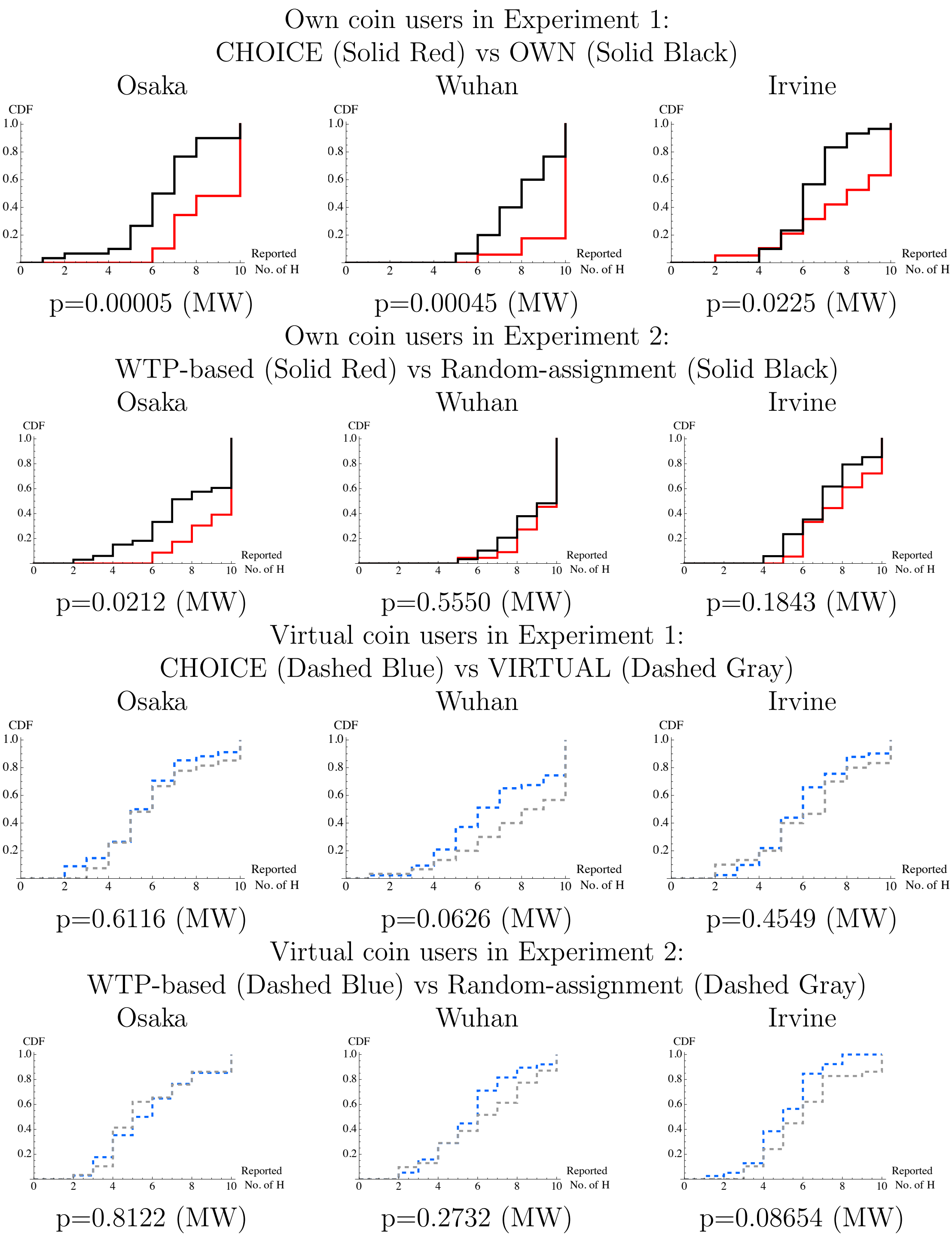


Figure E.3: CDFs of reported number of heads depending on the coin used and the way it is assigned. Own coin: WTP-based (Solid Red) vs random assignment (Solid Black). The virtual coin: WTP-based (Dashed Blue) vs random assignment (Dashed Gray). p-values are from Mann-Whitney test. These location-level comparisons are descriptive and are not adjusted for multiple comparisons.

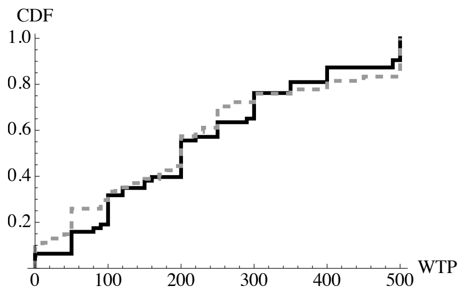
The main message is that across locations, while those who paid to use their own coin reported more heads than those who used their own coin for free (except in Wuhan and Irvine in Experiment 2), there are no significant differences in reported number of heads among those who used the virtual coin.

F Pilot session in Osaka

In Osaka, a pilot experiment of the CHOICE treatment in which everything was denominated in the local currency instead of points was conducted. Here we compare the submitted WTPs in the pilot session and the result of the CHOICE treatment in Experiment 1.

The pilot sessions (two sessions, 26 and 28 participants each) were conducted in March 2023. Participants were recruited from the same subject pool in Osaka, and the pilot sessions were conducted by the same author in the same manner (i.e., online using Zoom) as the main experiment. There were 30 males and 24 females, with an average age of 22.9. Participants earned on average 1180 JPY, including a 500 JPY show-up fee.

Figure F.1 compares the CDFs of submitted WTP in the pilot session and in the CHOICE treatment. There is no significant difference between the two (p=0.6193, Mann-Whitney test, two-tailed).



	Pilot	CHOICE	p-value
Mean	217.59	225.56	0.6193
(s.d.)	(167.51)	(150.48)	

p-value based on Mann-Whitney test, two-tailed.

Figure F.1: Cumulative distributions of willingness to pay to use own coin in CHOICE in Osaka (Solid Black) and Pilot in Osaka (Dashed Gray)

G Comparing heteroskedasticity-robust and cluster-robust standard errors

Tables G.1 to G.4 compare heteroskedasticity-robust (in parentheses) and cluster-robust (in square brackets) standard errors. As noted in footnote 17 of the main text, the latter tends to be smaller than the former. These estimates are reported for comparison. With as few as nine sessions in the Experiment 1 columns, cluster-robust standard errors are themselves unreliable; we report them to show that the choice of variance estimator does not drive any conclusion, not as a preferred alternative specification. The pattern of results is essentially unchanged. The one exception is the coefficient on WTP in Table G.3, which is significant at the 5% level under heteroskedasticity-robust standard errors but not under session clustering. The same coefficient remains significant under both estimators in Table G.4, so the selection result does not rest on the choice of variance estimator.

Table G.1: Comparing heteroskedasticity-robust and cluster-robust standard errors: Dependent variable is WTP

	Exp 1		Exp 2		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Wuhan	−65.722***	−61.625**	39.282*	44.855**	−65.722***	−52.264**
	(25.051)	(25.367)	(20.502)	(22.237)	(24.983)	(23.335)
	[24.989]	[15.207]	[22.477]	[23.369]	[24.005]	[18.458]
Irvine	−48.556*	−39.090	−2.718	24.581	−48.556*	−27.188
	(26.754)	(26.947)	(18.172)	(20.472)	(26.682)	(25.543)
	[25.431]	[25.527]	[24.979]	[21.643]	[24.430]	[21.984]
Exp 2					−49.421**	−46.891**
					(23.355)	(21.862)
					[30.686]	[26.084]
Exp 2 × Wuhan					105.004***	90.429***
					(32.336)	(29.656)
					[32.750]	[29.319]
Exp 2 × Irvine					45.838	46.455
					(32.296)	(30.665)
					[34.782]	[32.815]
Female		−5.315		−25.808		−18.063
		(19.919)		(16.155)		(12.540)
		[20.362]		[10.927]		[9.677]
Other		52.420		−59.384		−8.234
		(45.773)		(60.855)		(41.666)
		[56.077]		[62.471]		[44.284]
Age		0.570		−0.458		0.748
		(3.343)		(3.539)		(2.502)
		[4.587]		[3.239]		[2.912]
Risk		16.611***		6.386*		10.260***
		(4.449)		(3.594)		(2.773)
		[6.464]		[3.001]		[2.874]
Ethics		−3.539		2.438		−0.956
		(6.032)		(5.976)		(4.248)
		[4.922]		[5.268]		[3.828]
Gov. right		17.714		16.117		18.519**
		(15.001)		(11.593)		(9.024)
		[6.550]		[9.138]		[5.590]
No. Truthful		−5.563***		−3.708**		−4.495***
		(1.721)		(1.577)		(1.173)
		[1.994]		[1.229]		[1.002]
No. WTP>0		6.472***		9.078***		8.136***
		(1.909)		(1.481)		(1.151)
		[2.436]		[0.995]		[0.910]
Constant	225.556***	73.948	176.134***	6.827	225.556***	48.243
	(18.964)	(96.078)	(13.683)	(83.796)	(18.913)	(64.434)
	[24.340]	[93.463]	[20.049]	[83.382]	[23.382]	[66.197]
Adjusted R^2	0.028	0.210	0.011	0.150	0.014	0.167
N	183	180	359	359	542	539
No. of Clusters	9	9	17	17	26	26

Heteroskedasticity-robust standard errors are reported in parentheses, while cluster-robust standard errors at the session level are reported in square brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (based on heteroskedasticity-robust standard errors)

Table G.2: Comparing heteroskedasticity-robust and cluster-robust standard errors: Dependent variable is reported number of heads

	Exp 1 (CHOICE)		Exp 2 (WTP-based)		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Wuhan	1.014***	0.981**	-0.046	-0.271	1.030***	0.877**
	(0.374)	(0.381)	(0.356)	(0.404)	(0.370)	(0.348)
	[0.269]	[0.245]	[0.175]	[0.236]	[0.274]	[0.238]
Irvine	-0.073	-0.090	-0.834**	-0.561	-0.056	-0.053
	(0.380)	(0.403)	(0.355)	(0.394)	(0.373)	(0.372)
	[0.246]	[0.209]	[0.282]	[0.404]	[0.233]	[0.170]
Exp2					0.172	0.114
					(0.357)	(0.326)
					[0.174]	[0.138]
Wuhan×Exp2					-1.095**	-1.058**
					(0.512)	(0.474)
					[0.318]	[0.286]
Irvine×Exp2					-0.741	-0.489
					(0.506)	(0.485)
					[0.346]	[0.266]
WTP	0.005***	0.002	0.003**	0.003***	0.004***	0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]	[0.001]
Virtual coin	-0.943	-1.519**	-1.748***	-1.035	-1.270***	-1.195***
	(0.638)	(0.626)	(0.607)	(0.637)	(0.454)	(0.432)
	[0.683]	[0.660]	[0.493]	[0.627]	[0.440]	[0.412]
WTP × Virtual coin	-0.006**	-0.001	-0.006**	-0.008**	-0.006***	-0.005**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
	[0.002]	[0.002]	[0.002]	[0.003]	[0.001]	[0.002]
No. Truthful		-0.176***		-0.109***		-0.144***
		(0.028)		(0.031)		(0.020)
		[0.023]		[0.041]		[0.023]
No. WTP>0		-0.035		0.001		-0.023
		(0.035)		(0.033)		(0.024)
		[0.024]		[0.035]		[0.021]
Constant	6.891***	8.448***	7.939***	8.945***	7.227***	8.591***
	(0.514)	(1.624)	(0.565)	(2.006)	(0.400)	(1.223)
	[0.647]	[1.498]	[0.507]	[1.647]	[0.440]	[1.033]
Individual characteristics	No	Yes	No	Yes	No	Yes
Adjusted R^2	0.266	0.384	0.432	0.470	0.349	0.431
N	183	180	174	174	357	354
No. of Clusters	9	9	17	17	26	26

Heteroskedasticity-robust standard errors are reported in parentheses, while cluster-robust standard errors at the session level are reported in square brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (based on heteroskedasticity-robust standard errors)

Table G.3: Comparing heteroskedasticity-robust and cluster-robust standard errors: Selection (Dependent variable is the reported number of heads)

	(1)	(2)	(3)	(4)
WTP	0.002	0.003**	0.002	0.002**
	(0.0012)	(0.0013)	(0.0011)	(0.0012)
	[0.0011]	[0.0017]	[0.0009]	[0.0016]
virtual coin	-1.722***	-1.171**	-1.284***	-0.975*
	(0.322)	(0.547)	(0.321)	(0.509)
	[0.348]	[0.608]	[0.306]	[0.512]
WTP× virtual coin		-0.003		-0.002
		(0.002)		(0.002)
		[0.002]		[0.002]
Wuhan	0.792*	0.681	0.915*	0.843*
	(0.417)	(0.423)	(0.505)	(0.509)
	[0.220]	[0.196]	[0.293]	[0.287]
Irvine	-0.177	-0.246	0.179	0.113
	(0.404)	(0.397)	(0.470)	(0.468)
	[0.340]	[0.316]	[0.378]	[0.363]
No. Truthful			-0.184***	-0.181***
			(0.030)	(0.030)
			[0.022]	[0.021]
No. WTP>0			-0.022	-0.023
			(0.032)	(0.032)
			[0.035]	[0.036]
Constant	7.249***	7.043***	8.261***	8.192***
	(0.384)	(0.416)	(2.086)	(2.101)
	[0.369]	[0.457]	[2.701]	[2.700]
Individual characteristics	No	No	Yes	Yes
Adjusted R^2	0.164	0.168	0.283	0.282
N	185	185	185	185
No. of Clusters	17	17	17	17

Heteroskedasticity-robust standard errors are reported in parentheses, while cluster-robust standard errors at the session level are reported in square brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (based on heteroskedasticity-robust standard errors)

Table G.4: Comparing heteroskedasticity-robust and cluster-robust standard errors: Licensing (Dependent variable is the reported number of heads)

	Own coin users WTP-based vs Random-assignment		Own coin users who paid CHOICE and WTP-based pooled	
	(1)	(2)	(3)	(4)
WTP	0.003***	0.002**	0.003*	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
	[0.001]	[0.001]	[0.002]	[0.002]
WTP-based	0.352	0.373		
	(0.281)	(0.252)		
	[0.229]	[0.233]		
P			0.001	0.001
			(0.001)	(0.002)
			[0.001]	[0.001]
Wuhan	0.443	0.988**	0.545**	0.774***
	(0.330)	(0.385)	(0.273)	(0.282)
	[0.238]	[0.401]	[0.305]	[0.237]
Irvine	-0.831**	-0.083	-0.756*	-0.608
	(0.364)	(0.372)	(0.391)	(0.410)
	[0.354]	[0.266]	[0.314]	[0.291]
No. Truthful		-0.163***		-0.103***
		(0.027)		(0.029)
		[0.023]		[0.021]
No. WTP>0		0.027		0.022
		(0.034)		(0.044)
		[0.040]		[0.044]
Constant	7.310***	7.637***	7.563***	10.301***
	(0.393)	(1.701)	(0.391)	(1.551)
	[0.373]	[1.920]	[0.499]	[1.788]
Individual characteristics	No	Yes	No	Yes
Adjusted R^2	0.190	0.369	0.180	0.291
N	159	159	128	126
No. of Clusters	17	17	23	23

Heteroskedasticity-robust standard errors are reported in parentheses, while cluster-robust standard errors at the session level are reported in square brackets.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (based on heteroskedasticity-robust standard errors)