Stochastic Dominance and Demand for Surprise: Supplementary information

Séverine Toussaert*

March 26, 2024

Abstract

This appendix provides additional background information about experimental procedures, pilot data, variables collected, descriptive results, and theoretical analyses, which could not be reported in the main appendix due to space restrictions.

*All remaining errors are mine. Contact: `severine.toussaert@economics.ox.ac.uk` for any questions about the project.

Contents

A Experimental procedures and sample **2**
A.1 Preliminary pilot . . . 2
A.2 Data collection process and final sample . . . 5
A.3 Prize allocation and distribution . . . 9
A.4 Changes in preferences with the spread of the pandemic . . . 10

B Preferences $\succeq$ on X and implied SD violations **20**
B.1 Incentive compatibility of the elicitation procedure . . . 20
B.2 Data on ranking and valuation of the destinations . . . 24
B.3 Preference consistency and prevalence of SD violations . . . 25
B.4 More information on the size and shape of SD violations . . . 26

C Decisions in the design task **28**
C.1 Description of lottery-building exercise . . . 28
C.2 Characteristics of the favorite option . . . 29
C.2.1 Choice of destination(s) and probability weights . . . 29
C.2.2 Preferences for delaying the resolution of uncertainty . . . 32
C.3 Preferences in the binary choice exercise vs. design task . . . 34
C.3.1 Link between SD violations and choice to design a lottery . . . 34
C.3.2 Link between SD violations and preferences for delay . . . 35

D Drivers of preferences for surprise trips **36**
D.1 Psychological motives behind wildcard valuations . . . 36
D.2 Attitudes towards surprise trips . . . 39

E Knowledge and preferences over trip attributes **41**
E.1 Knowledge of trip destinations . . . 41
E.2 Preferences over trip attributes . . . 42

F Alternative mechanisms **44**
F.1 Ambiguity aversion and hedging . . . 44
F.1.1 Ambiguity aversion in the domain of money . . . 44
F.1.2 Hedging due to mistrust . . . 48
F.2 Non-linear probability weighting . . . 49

A Experimental procedures and sample

A.1 Preliminary pilot

A pilot was run in December 2019 with a small sample of UK travelers recruited on social media (more on recruitment in Subsection A.2). The goal of this pilot was threefold: (i) test whether the recruitment strategy and logistics were working; (ii) assess whether participants understood the core tasks; (iii) sharpen the design by improving the selection of lotteries and measurement of preferences. Given the peculiarity of the context, the pilot explored a wider range of decision tasks in less depth to understand possible sources of variation. Below I explain the contents of the pilot survey (see summary in Table A1) and how it informed the main study design.[1]

Type of lotteries The pilot considered not only risky lotteries but also lotteries with unknown probability distribution and/or unknown support in order to assess the impact of ambiguity and unawareness on preferences for randomization:

1. In 18 binary decision problems (DPs), participants chose between a sure destination and a risky lottery.

2. In addition, 12 DPs presented binary-outcome lotteries with unknown probabilities and/or unknown support. Some problems involved a trip to a new and unknown destination either as a single outcome or as part of a lottery.

Cardinal information The survey not only elicited valuations of the 10 destinations (Part 1), but also of a subset of lotteries (+ wildcard trip) (Part 4). Valuations were elicited twice in a row for each option i.e., with the destination revealed either (i) no earlier than a week before travel ("Discover later") or (ii) at the end of the survey ("Discover now").

Timing of information revelation To assess how the date at which the destination is revealed might affect preferences for randomization, respondents were asked in Part 3 to consider 14 DPs in which the chosen destination in Option B would be disclosed right at the end of the survey (labelled as "no surprise" lotteries). The

[1]The online questionnaire for this pilot is available **here**.

Table A1: Structure of the pilot survey

PART 1	**Elicitation of preferences $\succeq$ over holiday trips** ▹ Ordinal ranking of 10 destinations, $x_1, x_2, ..., x_{10}$ ▹ Monetary valuation $v_k \in [0, 500]$ for each destination x_k
PART 2	**18 binary decision problems (DPs) with risky "surprise" lotteries** A: $(x_i, 1)$ vs. B: $(x_j, p_j; x_k, p_k; ...; x_l, p_l)$ where B is either a dominated lottery (13 DPs), a dominant one (3), or involves no dominance relation (2). Destination revealed $\leq$ 1 week before travel.
	12 DPs with ambiguous "surprise" lotteries and/or wildcard trip x_* A: $(x_i, 1)$ vs. B: x_* or $(x_j, ?; x_k, ?)$ where $x_k \in \{x_1, ..., x_{10}, x_*\}$ where B is either a dominated lottery (3 DPs), a dominant one (2), or involves no dominance relation / contains x_* (7). Destination revealed $\leq$ 1 week before travel.
PART 3	**14 DPs with "no surprise" lotteries** Subset of DPs presented in SECTION 2 but with the destination revealed at the end of the survey (10 with risky lotteries, 4 with ambiguous ones and/or wildcard trip).
PART 4	**Valuation of surprise trips** Monetary valuation of the wildcard trip and a subset of lotteries from SECTIONS 2 and 3.
	Design of favorite lottery ▹ Selection of support in $\{x_1, ..., x_{10}, x_*\}$ and probability distribution $\mathbf{p}$ or **?** (unknown) ▹ Choice to reveal destination either at the end of the survey or $\leq$ 1 week before travel.
PART 5	**Preferences over monetary gambles** ▹ Design favorite lottery as in SECTION 4 but over gift card amounts chosen from the set $\{420, 400, 350, 300, 250, 200, 150, 100, 50, 20, m_*\}$, where m_* is unknown. ▹ Valuation of one risky and one ambiguous bet
	Questions about travel history and preferences
END	Selection of choice problem

chosen DPs were a subset of the problems presented in PART 2, to enable direct comparisons.

Lottery-building exercise Respondents performed two lottery-building exercises, one with trip destinations (PART 4) and the other with gift card amounts (PART 5). In both cases, the set of options included a wildcard (unknown destination or monetary amount) as well as 10 known options (the 10 destinations already presented or 10 gift cards amounts). Participants first picked the support (which could be degenerate) among the set of 11 options and then decided whether to assign specific chances to each selected option (choice of given **p**) or leave it to chance (unknown probability **?**). Finally, they chose whether to learn the outcome of the lottery at the end of the survey ("Discover now") or around the time of their preferred travel date ("Discover later").

Other differences Some other minor differences included:

- List of destinations: All destinations were the same at the exception of Lake Garda (Italy), which was replaced in the main study by Porto (Portugal). This change was made due to the early spread of COVID-19 in Italy and the main data collection happening at the end of February / beginning of March. In addition, the pictures of the destinations were different.

- Design of favorite lottery: Although respondents in the pilot were allowed to design a degenerate lottery as their favorite option, they might have been confused by this framing and not realized the possibility of selecting a single destination. In the main study, respondents were therefore explicitly offered to either select a sure destination or build their favorite lottery with at least two outcomes.

- Wildcard trip: The trip to a new and unknown destination was initially called "Joker trip" instead of "wildcard trip". The name was changed in order to avoid any negative connotation that could lead respondents to think that they would be tricked if they chose this option.

- Reasons for randomization: The end of the pilot survey asked questions about information preferences more generally / in other domains. Instead, the main survey strictly focused on attitudes towards surprise trips in order to better understand the specific context.

Main takeaways This small pilot produced the following takeaways:

1. Strict focus on risky lotteries: The number of different tasks was too large, with respondents having to keep track of too many dimensions. In the main survey, I decided to strictly focus on risky lotteries, for which I had a clear theoretical framework in mind; on the other hand, the modeling of attitudes towards

ambiguity and unawareness is much less obvious, if only because the quantification of uncertainty in an ambiguous setting strongly depends on the model of ambiguity assumed. The set of DPs with risky lotteries was expanded from 18 to 39 in the main study, and the 12 DPs with ambiguous lotteries and/or wildcard trip were removed.[2] In addition, I included 6 DPs in which Option B is a degenerate lottery so as to test for the robustness of the preferences over destinations measured via the ranking method.

2. Removal of cardinal information on lotteries: The monetary valuations of lotteries in Part 4 displayed limited variation across lottery types (including delay); in addition, it was unclear whether levels could be trusted given the general noisiness of WTP/WTA data, making inferences about preferences for surprise trips vs. sure destinations difficult. As a result, the new survey only measured respondents' valuations for (i) the 10 destinations, to quantify the size of SD violations and measure indifferences; (ii) the wildcard trip, to study whether SD violations were related to preferences for this unknown destination.

3. Measurement of preferred delay: There was little variation in decisions between Parts 2 & 3, meaning that randomization behavior did not appear to shift depending on whether the destination was revealed at the end of the survey or close to the time of travel (in line with the previous point). While this suggests that substantive delay of the resolution of uncertainty might not be a necessary condition for randomization, the chosen setup made it difficult to assess whether respondents actually value any delay. To better understand preferences for delay, I instead allowed respondents in the main survey to directly select their preferred revelation date when designing their favorite lottery and measured their WTP for this option; the binary problems with no delay were removed.

A.2 Data collection process and final sample

Participant pool The pilot described in Subsection A.1 was completed by 9 London-based travelers recruited via Facebook ads run by the company. The ad only targeted respondents based on their location and age i.e., living in London and being at least 18 years old. Importantly, the name of the company did not appear on the ad in order

[2]Most of 18 DPs presented in the pilot were kept in the set, with some minor modifications to make the design more symmetric.

Figure A1: Facebook ad

Notes: Facebook ad used to recruit participants for the pilot conducted in December 2019. A similar ad was used in February/March 2020 for the main recruitment with a slight variation in wording and a different picture.

to avoid recruiting travelers specifically interested in surprise holiday experiences (see Figure A1 for a screenshot). A link to the consent form was included in the ad for participants to sign up.

For the main study, the company ran Facebook ads from 27 February 2020 until around 10 March 2020 (similar announcement as in December 2019). Due to the low number of sign-ups, the recruitment strategy was adjusted and participants were instead recruited via the Behavioural Research Lab of the London School of Economics and Political Science (https://www.lse.ac.uk/management/research/bl). Recruitment via the LSE Lab had two main advantages: an easy targeting of London-based travelers and access to a fairly diverse pool of student and non-student participants (relative to some other labs). A recruitment email was sent by the LSE Lab manager on 9 March 2020 with the following announcement: "*Participate in a holiday trip* *study conducted by a researcher at the University of Oxford for the chance to win a free holiday trip! Very flexible booking and travel dates. To learn more about the research study and sign up to participate, click* **here**".

Onboarding procedure, attrition and final sample The collection of survey responses proceeded as follows:

1. Interested people signed up for the study by completing the consent form and providing an email address. Enrollment was on a first-come, first-served basis and capped at 20 respondents for the pilot and 100 respondents for the main study. Recruitment was closed once the cap was reached. Respondents in the main study were made aware of this cap (this was not the case for the pilot).

2. Respondents who consented to participate were emailed the survey link within 24h hours of signing up, together with a unique participant ID. The invitation email to complete the main study is shown in Figure A2.

3. Respondents were originally given 7 days to complete the survey, although the deadline was extended by a few days for the main study due to the extraordinary circumstances around COVID-19. They were sent two reminders.

4. Once respondents clicked on the link, they were instructed to complete the survey within 2 hours. At several time points during the survey, a timer showing the remaining time was displayed at the top right-hand corner of the screen. Respondents were prevented from proceeding if they hit the mark of 7200 seconds (2 hours) at the beginning of PARTS 2, 3, 4 and 5 (blocks "Monetary lotteries" and "You as a traveller"). No timer was applied upon disclosing prizes, meaning that total survey time could in principle exceed 7200 seconds.

Final sample Among the 20 respondents who signed up for the pilot, 9 provided a complete response.[3] For the main study, 83 of the 100 people who originally consented to participate finished the survey. Among the 17 respondents who did not complete the study, 13 left the survey by themselves (all with $\leq$ 20 % completed) and 4 were timed out (3 of them before completing Section 2).

[3]Two additional respondents opened the survey and one of them completed about half of it.

Figure A2: Invitation email to complete the survey

Holiday Trip Study <holidaytripstudy@gmail.com> Thu, Mar 5, 2020, 12:26 PM

to me

Dear curious traveller,

Thank you so much for agreeing to participate in our Holiday Trip Study! We are really excited to learn more about your preferences for holiday trips!

We capped this study at **100 participants**. If you believe that you are no longer able to participate, please let us know ASAP so we can give your spot to another person.

We have **5 trips to a European destination to give away (£420 value).** Booking and travel dates are fully flexible (until 31 Dec 2021).

Other prizes include Amazon gift cards or equivalent cash amounts (one large prize, 4 smaller prizes) and trip vouchers. In addition, **all participants will get a £25 voucher valid for a trip with our travel partner.**

You will have **7 days to complete the survey** and we estimate that it should take you **about 30-45 minutes**.

Note: You should **take no more than 2 hours to complete it** once you click on the survey link. After that, the survey will stop being accessible.

To get started, you will need to enter a Participant ID number. Your ID number is:

XXXXX

Please keep this information for your records. If you are ready, click below:

Take the survey

Once again, thank you so much for your participation! Your feedback is really important to us.

If you have any questions, do not hesitate to respond to this email.

The Holiday Trip Study Team

A.3 Prize allocation and distribution

For the initial pilot, no information was provided about the targeted number of participants or the number of prizes. In the main study, respondents were told both in the consent form (link **here**) and subsequent invitation email (see Figure A2) that participation was capped at 100 respondents. In addition, the invitation email specified that we could give away 5 trips to a European destination worth £420, implying a 1 in 20 chance of winning a trip. The survey itself did not specify the specific chances. Prizes were allocated based on a participant ID, which was randomly determined prior to the start of the study. The choice to realize the uncertainty about the prize allocation prior to taking the survey was made so respondents do not believe that their answers can influence what section/task is chosen to determine their prize; in addition, this made it logistically easy to reveal the selected prize right at the end of the survey, thus ensuring that this uncertainty was short-lived.

Below I discuss how the prize allocation was made. First, I chose how to allocate the chances of receiving a prize based on the various survey parts (called "sections" in the survey), with more trips being allocated to parts that presented a larger number of decisions (PARTS 1 and 2); for respondents who received a trip based on PART 2, I randomly drew one decision problem that applied to all of them (DP #7). To minimize programming costs, I drew at random which of PART 3 and 4 would be used to allocate the remaining number of trips. The selection was done using Stata and the code generated for the prize allocation is available on OSF.[4]

In total, 4 participants were awarded a trip based on PART 2 (DP #7), two participants received a random (BDM) payment based on PART 1, one participant received £350 based on PART 5 (outcome of monetary lottery) and 3 participants received a payment based on a valuation task for an ambiguous £50 bet; the remaining participants received trip vouchers valid for the destination preferred in DP #7 of PART 2. The prize breakdown is shown in A2.

Due to the COVID-19 crisis, respondents received their trip prizes with substantial delay. Although voucher codes were sent by the company in spring 2020, the business was shut down the following year, implying that respondents were largely unable to benefit from the offer given the travel restrictions in place. Identifying a travel partner who could deliver the trips promised at a cost of £420 each was no easy task.

[4]Note that some of the randomization (e.g., selection of BDM payments) was done within the survey.

Table A2: Prize distribution

Part*	Prize type	Number offered	Monetary value
1	random payment	2	£499, £207 (BDM draw)
2	trip	4	£420
2	trip voucher	72	£20
2	trip voucher	1	£50
5	monetary lottery payment	1	£350
5	bet on ambiguous box	3	£50 (×2), £49 (BDM draw)

Notes: *Called "Section" in the actual survey.

However, a solution was found in summer 2022 by hiring a individual travel agent via Travel Counsellors (https://www.travelcounsellors.co.uk/). An announcement with booking details was sent to participants in September 2022, with a booking deadline of 31 December 2023. With this new solution, 3 of the 4 trips were booked (the email of one of the winners expired).

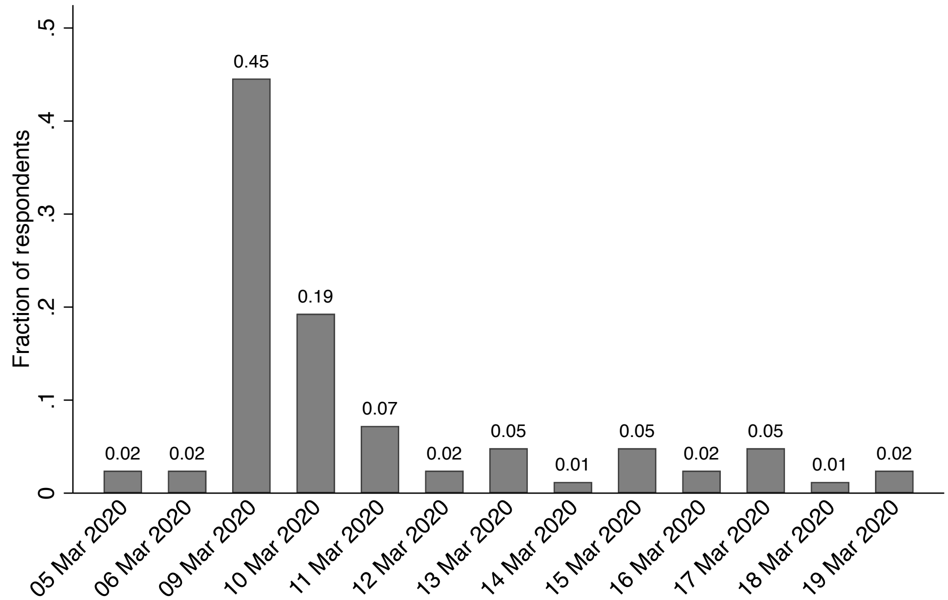
A.4 Changes in preferences with the spread of the pandemic

Below I provide a more extensive discussion of the impact of the onset of the pandemic on survey responses. The data collection took place during the month of March 2020, at a time of growing travel restrictions due to the spread of the COVID-19 pandemic. An important question is whether preferences for travel and surprise holidays shifted in significant ways as a result. While I do not have a source of exogenous variation to give a causal answer, below I exploit natural variation in the date at which the survey was completed in order to shed some light on the possible impact of COVID-19. On 9 March 2020, Italy was the first European (and worldwide) country to impose a national lockdown, an event which made the consequences of the spread of the pandemic particularly salient. Coincidentally, it was on this same day that participants from the LSE lab pool were emailed about the study, generating a spike in survey completion (see Figure A3).

Below I examine if/how respondents' decisions differ based on whether they completed the survey after the lockdown announcement of March 9th (i.e., from March 10th onwards, N = 42) or up to that date (N = 41).[5] A summary of mean differences

[5]The official announcement of a nationwide lockdown was made by Giuseppe Conte, then Italian prime minister, at a press conference late on Monday 9 March 2020 (https://www.theguardian.com/

Figure A3: Distribution of survey completion dates



for a large range of outcome variables is presented in Table A3; for key variables, I also discuss potential distributional changes.

Finding 1 (*Shift in preferred travel date*) The first finding is that respondents post-lockdown announcement shifted their preferred travel date by 105 days on average, a large shift relative to a baseline date of 176 days into the future (shift from Aug/Sept 2020 to Dec 2020). Figure A4 shows a clear first-order stochastic dominance shift in the distribution of chosen travel dates following the announcement. Despite this, the majority of respondents (85% pre- and 65% post-lockdown announcement) still hoped to be able to travel by December 2020 (see Figure A5).

Finding 2 (*Stronger preference for delay*) The second finding is that among the 46 respondents (out of 83) who selected a trip lottery as their favorite option, those who answered the survey post-announcement had a stronger preference for delaying the resolution of uncertainty about the destination. First, they were more likely to choose a revelation date that is not "*Today, after I completed the survey*", thus postponing by at least one day (87% vs. 61%, $p < 0.05$). Differences between the two groups of respondents also persist conditional on preferring any delay. In particular, the length

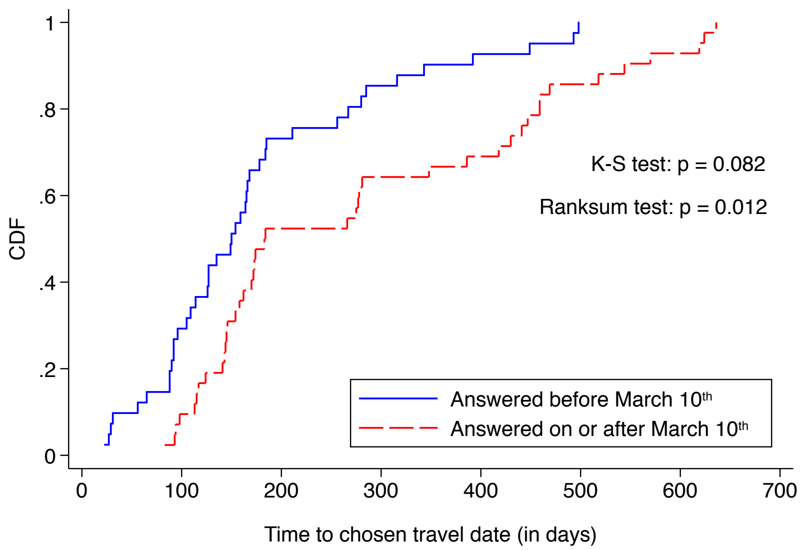
`world/2020/mar/09/coronavirus-italy-prime-minister-country-lockdown`).

Table A3: Preferences pre- and post-lockdown announcement (March 9th)

	Pre-	Post-	Diff.	s.e.	N
SD violations in binary choices					
Number of violations on $\mathcal{D}^+$	4.0	4.8	-0.8	(1.1)	83
Fraction of violations on $\mathcal{D}^+$	0.18	0.27	-0.09*	(0.05)	79
Number of violations on $\mathcal{B}^+$	2.3	2.4	-0.1	(0.6)	83
Fraction of violations on $\mathcal{B}^+$	0.21	0.30	-0.09	(0.06)	79
Number of violations on $\mathcal{B}^-$	1.4	1.8	-0.4	(0.6)	83
Fraction of violations on $\mathcal{B}^-$	0.13	0.19	-0.06	(0.06)	79
Favorite option					
Sure destination	0.44	0.45	-0.01	(0.11)	83
Trip lottery with fixed delay	0.39	0.40	-0.01	(0.11)	83
Trip lottery with custom delay	0.17	0.14	0.03	(0.08)	83
Valuation of wildcard trip (in £)	266.7	283.9	-17.3	(29.6)	83
Chose to reveal the destination later	0.61	0.87	-0.26**	(0.13)	46
Length of delay (in days)	29.1	109.6	-80.5***	(26.7)	46
Time to chosen travel date (in days)	176.4	281.7	-105.3***	(33.0)	83
% delay relative to travel date	23.4	56.6	-33.2**	(12.4)	46
WTP for delay (in £)	5.2	18.0	-12.8***	(4.0)	46
Surprise trip like rating (0-100)	52.9	53.6	-0.7	(6.8)	83
Total survey time (in minutes)	43.1	45.0	-1.9	(4.4)	83

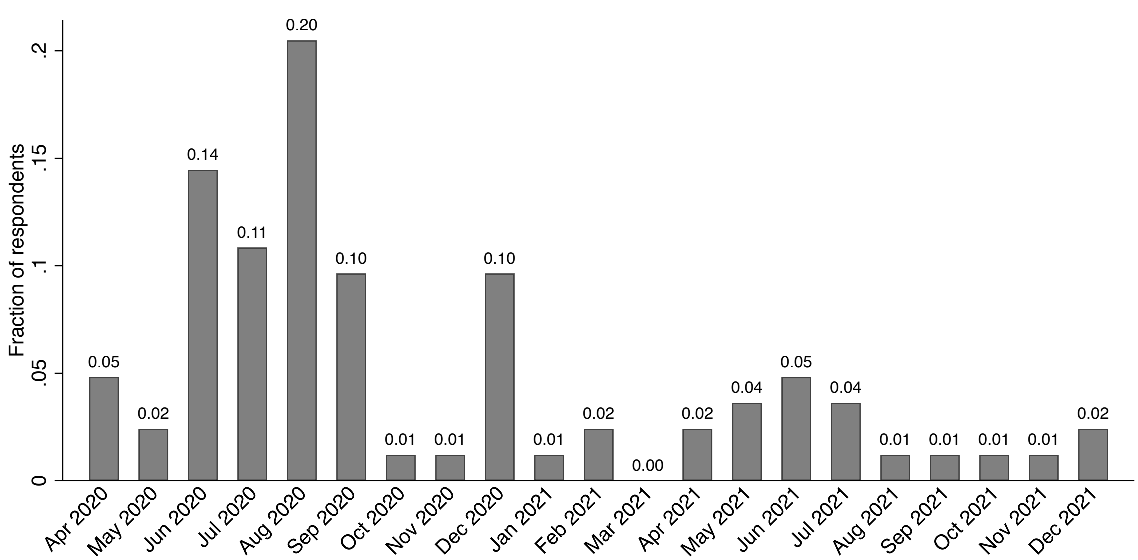
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. `Chose to reveal the destination later` is an indicator equal to 1 if a respondent chose not to reveal the trip destination of their favorite lottery right away; `Length of delay (in days)` is the number of days the respondent was willing to postpone and `% delay relative to travel date` is the amount of delay as a percentage of `Time to chosen travel date (in days)`. The variable `WTP for delay (in £)` is the amount of money a respondent was willing to forego to delay learning about the destination of their favorite lottery; it is set equal to 0 for respondents who preferred not to delay the resolution of uncertainty in the first place (diff. $= -12.2$, s.e. $= 5.0$, p$= 0.02$ among the 34 respondents who preferred to delay and thus completed the MPL to measure WTP). Results qualitatively unchanged when using a Wilcoxon rank-sum test.

Figure A4: Chosen travel date pre- and post-lockdown announcement



Notes: p-values are from Kolmogorov-Smirnov and Wilcoxon rank-sum tests of equality of distributions. N = 83.

Figure A5: Distribution of chosen travel dates



of the chosen delay relative to the planned travel date is higher by 33 percentage points ($p < 0.05$) on average for post-announcement respondents. Furthermore, those who responded after March 9th generally had a higher WTP for delay (Figure A6), with an average WTP difference of £12.80 when including those who preferred not to delay in the first place ($N = 46$, $p < 0.01$) and an average difference of £12.20 among only those who completed the MPL exercise ($N = 34$, $p = 0.02$).

Finding 3 (*No significant change in preferences for randomization*) The third finding is that preferences for randomization appear to have changed very little following the announcement. As shown in Table A3, the prevalence of stochastic dominance violations does not significantly differ pre- vs. post-announcement, whether in the direction of a preference for randomization ($\mathcal{D}^+$ and $\mathcal{B}^+$ problems) or in the direction of a preference for certainty ($\mathcal{B}^-$ problems).[6] The two categories of respondents have also very similar preferences when designing their favorite option and valuing the wildcard trip.[7] Finally, the mean rating on the question "Overall, how much did you like the concept of "surprise trip" presented in this study?" is nearly identical pre- and post-announcement at around 53/100.

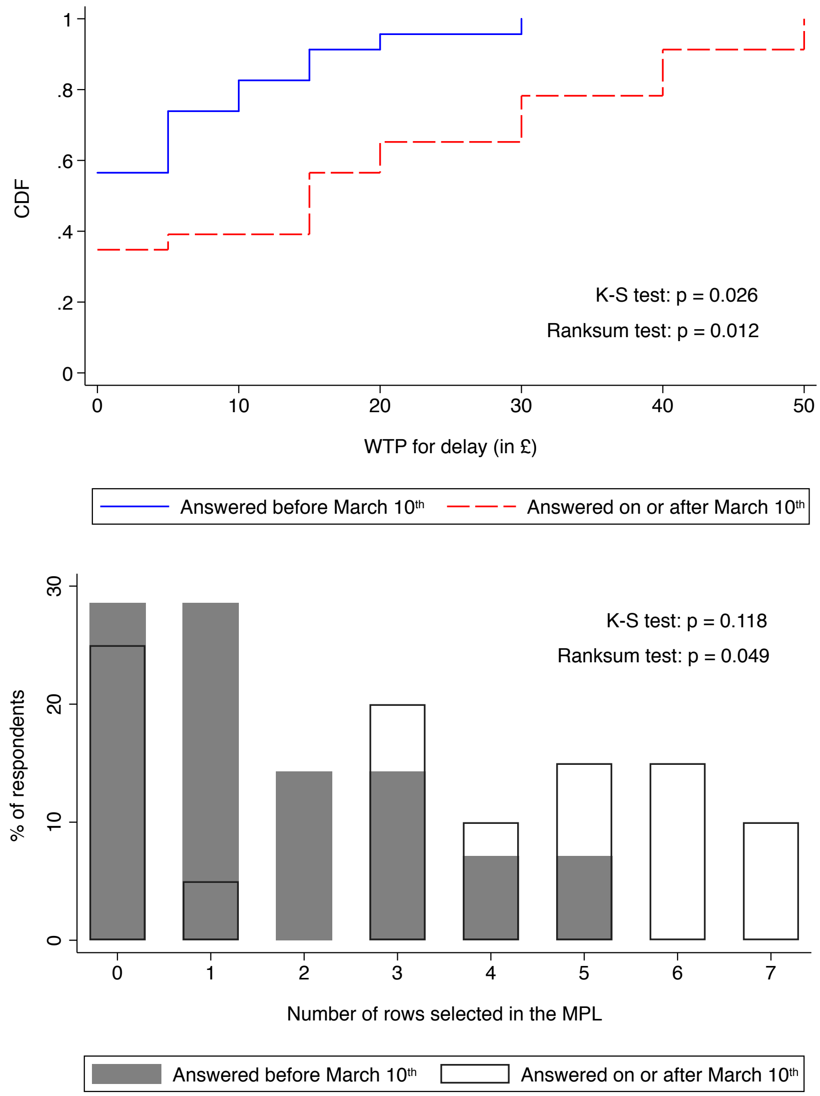
Since respondents pre- and post-lockdown announcement might differ in other ways than their survey completion date, I examine in a regression model whether the patterns observed in Table A3 persist after controlling for respondents' characteristics, finding that results are qualitatively unchanged.[8] In Table A4, I focus on preferences for the travel date ($N = 83$) and day at which the destination is revealed ($N = 46$), variables for which a statistically-significant relationship was found with survey completion date on raw comparisons. Column (2) shows that a large difference in the chosen travel date persists after controlling for observable respondent characteristics, with the date being pushed back by an average of 102.6 days post-lockdown announcement ($p = 0.005$). Since the length of the anticipation period depends on the chosen travel date, one may ask whether the reason why preferences for keeping the surprise of the destination appear stronger post-announcement is because travel

[6]However, one should note a marginally significant increase by 9.4 percentage points in the proportion of stochastic dominance violations on $\mathcal{D}^+$ problems.

[7]Among those who built a surprise lottery in the design task, features of the chosen lottery (support size, entropy of the distribution, expected value relative to the top option) are also distributed in a similar way whether they replied pre- or post-announcement.

[8]One small change is that the marginally significant difference in Table A3 in the proportion of SD violations on $\mathcal{D}^+$ becomes significant at $p < 0.05$ after controlling for respondents' characteristics.

Figure A6: WTP for delay pre- and post-lockdown accouncement



Notes: p-values are from Kolmogorov-Smirnov and Wilcoxon rank-sum tests of equality of distributions. The variable `WTP for delay (in £)` is the amount of money a respondent was willing to forego to delay learning the destination of their favorite lottery (N = 46); it is set equal to 0 for the 12 respondents who preferred not to delay the resolution of uncertainty in the first place (i.e., those who selected "*Today, after I completed the survey*" as their preferred revelation date). The variable `Number of rows selected in the MPL` corresponds to the number of times the respondent preferred "*Revealing Later*" over "*Revealing Now* + £X" in the MPL question; this question was only asked to those who indicated a preference for delay (N = 34).

dates were pushed back. Columns (3)-(8) in Table A4 show that, in fact, this does not seem to be the case. First, post-announcement respondents are over 30 percentage points more likely to prefer any delay in the revelation date even after controlling for their chosen date of travel; this holds with and without additional controls (columns (3) and (4)) and the effect is in fact larger than the raw difference of 26 percentage points in Table A3. Similarly, the length of the chosen delay remains large at about 75 days with all controls included ($p = 0.012$), and 87 days after only controlling for date of travel. Finally, column (5) shows that the average difference in WTP for delay remains large at £13.34 after controlling for the travel date ($p = 0.004$). On the other hand, the chosen delay length itself captures about half of the post-announcement effect on WTP for delay, with each additional day of postponing the revelation date being associated with an average increase of £0.09 with all controls (£0.08 without additional controls).

Finding 4 (*No change in trip valuations*) The fourth finding is that trip valuations did not go down following the announcement of the Italian lockdown. This is reassuring because one could have expected the uncertainty around future travel opportunities and the heightened exposure risks due to travel to greatly reduce the attractiveness of the trip packages offered in the study. Figure A7 shows that this does not appear to be the case. In fact, trip valuations pre- and post-lockdown announcement are nearly identically distributed for most destinations. The mean valuation for the destination ranked #1 is £376.34 pre-announcement and £370.10 post-announcement, and all respondents value their top destination at £200 or more at the exception of two people (both in the post-announcement group). In other words, the trips offered in the study remained an attractive prize even as the COVID-19 restrictions expanded. Related to this finding, I also find no statistically significant difference in people's preferences on various holiday trip criteria (e.g., quietness vs. vibrant atmosphere), suggesting that any preference shift that might have occurred during the pandemic appeared only until later (see Figure A8).

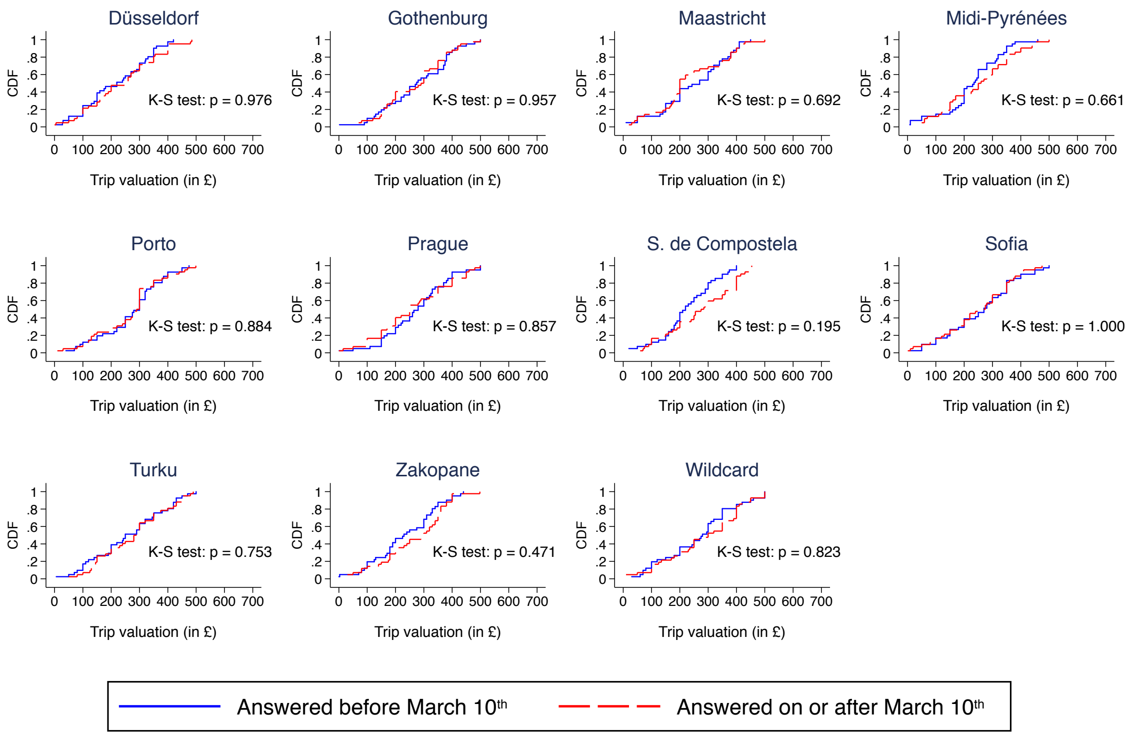
Summary Overall, preferences over holidays and surprise destinations appear to have remained stable over the time window of the survey (5 to 19 March 2020), with two exceptions: (i) respondents pushed back their anticipated travel date following the announcement of the Italian lockdown; (ii) their preference for preserving the surprise of the destination appears to have been stronger post-announcement. Two

Table A4: Preferences for delay and pre- and post-lockdown announcement

	(1) Travel date	(2) Travel date	(3) Chose to delay	(4) Chose to delay	(5) Delay length	(6) Delay length	(7) WTP for delay	(8) WTP for delay	(9) WTP for delay
Survey completed after March 9 lockdown	105.300***	102.552***	0.369***	0.333**	87.270***	74.960**	13.343***	6.802	6.579
	(32.968)	(35.745)	(0.132)	(0.139)	(29.187)	(28.278)	(4.408)	(4.256)	(4.325)
Male respondent		39.113		-0.115		-36.829			1.321
		(35.650)		(0.141)		(28.725)			(4.109)
Age		-1.496		0.002		-1.620			0.238
		(2.660)		(0.013)		(2.595)			(0.365)
Number of trips abroad in last 12 months		-3.113		-0.006		10.340**			-1.003
		(6.800)		(0.024)		(4.843)			(0.719)
Number of EU countries visited in list of 10		2.931		0.054		2.702			-0.663
		(9.754)		(0.039)		(8.013)			(1.123)
Traveling with one partner		1.695		0.253		-56.799			0.440
		(54.071)		(0.273)		(55.578)			(7.887)
Traveling with two partners or more		96.159		0.190		-72.884			-6.463
		(74.077)		(0.345)		(70.160)			(9.961)
Recruited via the LSE Lab		1.497		-0.018		-37.572			-5.636
		(50.278)		(0.221)		(45.006)			(6.357)
Travel date			-0.001**	-0.001	-0.056	-0.020	-0.004	-0.000	0.007
			(0.000)	(0.000)	(0.094)	(0.095)	(0.014)	(0.012)	(0.013)
Delay length								0.075***	0.094***
								(0.020)	(0.023)
Constant	176.390***	179.326*	0.739***	0.283	37.313	128.845	5.840	3.044	6.397
	(23.452)	(107.301)	(0.106)	(0.608)	(23.440)	(123.659)	(3.540)	(3.200)	(17.558)
Observations	83	83	46	46	46	46	46	46	46
Adjusted R^2	0.101	0.097	0.134	0.118	0.140	0.255	0.152	0.346	0.370

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

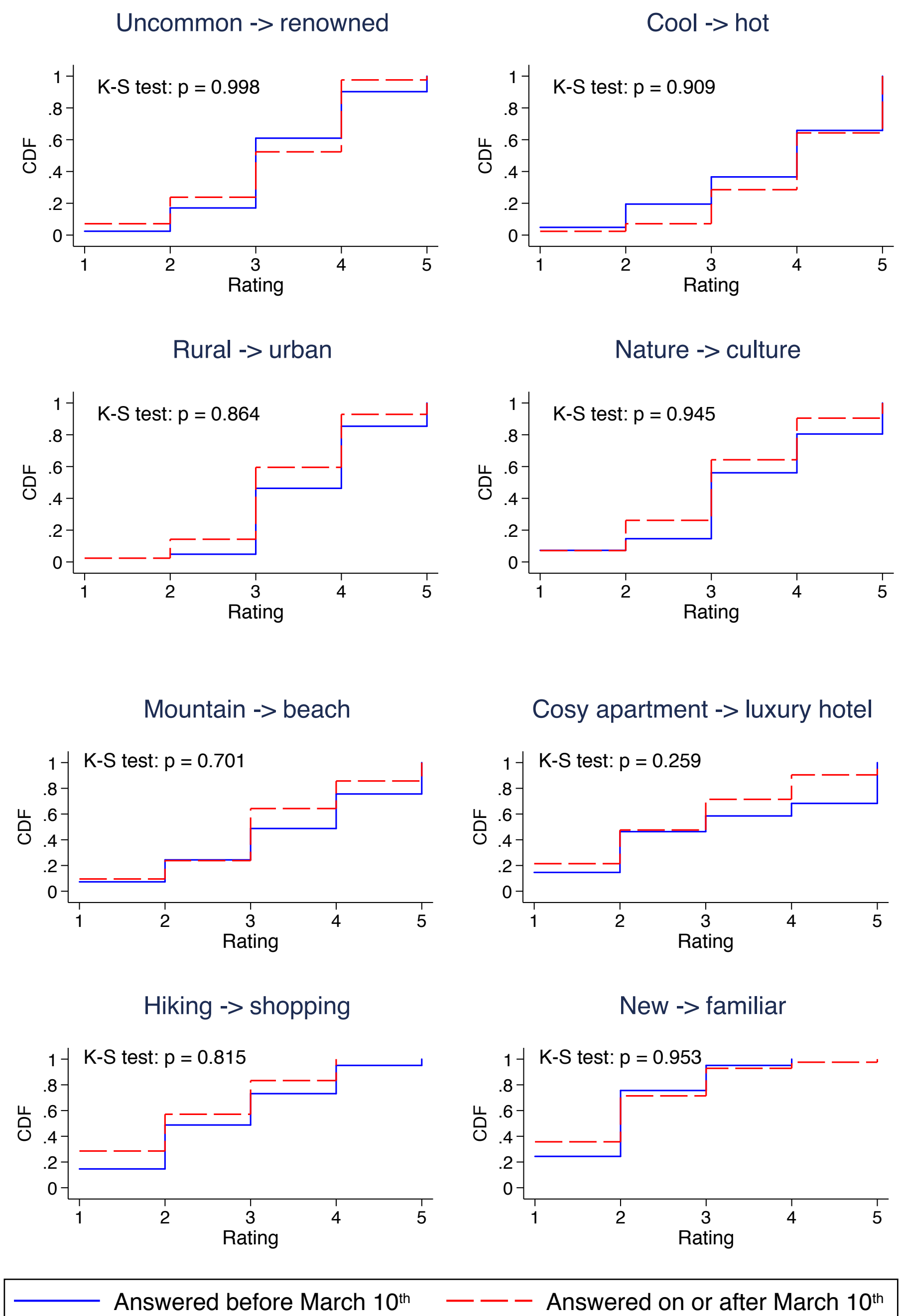
Figure A7: Trip valuations pre- and post-lockdown announcement



Notes: p-values are from Kolmogorov-Smirnov tests of equality of distributions.

important caveats to these results should be noted. First, respondents pre- and post-lockdown might differ on many unobservable characteristics that could explain the observed differences, meaning that no causal statement can be made. Second, the statistical analyses should be interpreted with caution due to the small sample sizes on which they were performed. For these two reasons, the results presented in this section only provide suggestive evidence of what might have been the effect of the onset of the pandemic on travelers' preferences.

Figure A8: Holiday trip preferences pre- and post-lockdown announcement



Notes: p-values are from Kolmogorov-Smirnov tests of equality of distributions.

B Preferences $\succeq$ on X and implied SD violations

B.1 Incentive compatibility of the elicitation procedure

Below I provide a more detailed examination of the incentive compatibility of the elicitation procedure adopted in this paper and report on empirical tests to assess evidence of bias. The survey elicited preferences over destinations in two steps: 1) ranking of the 10 trips listed in a random order; 2) elicitation of the minimum price $v \in [0, 500]$ for giving up each trip. To incentivize truthful reporting, I used the following procedure:

1. The chances of receiving a given trip (or equivalent amount of money) were higher if the trip was listed higher in the ranking. Formally, $p_k > p_{k'}$ for all $k' > k$, where p_k is the probability of implementation of the valuation question for x_k (option ranked $\#k$). The exact probability distribution was $\mathbf{p} = (0.25, 0.20, 0.15, 0.12, 0.10, 0.08, 0.05, 0.03, 0.02, 0)$; this information was available to the participant by clicking on a link.

2. Valuations are elicited using the BDM mechanism (Becker-DeGroot-Marschak, 1964). If x_k is selected in Step 1, the respondent's valuation v_k for that trip is compared to a random number $V \sim \mathcal{U}_{[0,500]}$. If $V \geq v_k$, the respondent receives £V; if $V < v_k$, the respondent receives the trip.

Incentive compatibility of Step 1 Incentive compatibility requires that the respondent satisfies stochastic dominance over the subset of lotteries having the same probability distribution $\mathbf{p} := (p_1, p_2, .., p_{10})$ (but a different support). This is the case for most measures of uncertainty considered in this paper e.g., $\psi(-\sum_{k=1}^{n} p_k \ln(p_k))$, $\psi(\sum_{k=1}^{n} p_k(1 - p_k))$, or $\psi(|\text{supp}(\mathbf{p})| - 1)$ where $\psi'(.) > 0$. This property is however not satisfied e.g., if the DM cares about the variance in valuations; in this case, the DM has an incentive to give a higher rank to their least preferred option.[9]

[9]To take a simple example, suppose the DM is asked to order 3 trips x, y, z and let $r_w \in \{1, 2, 3\}$ be the rank number they assign to $w \in \{x, y, z\}$. Assume their true preference ordering is $x \succ y \succ z$ with corresponding valuations $(v_x, v_y, v_z) = (450, 150, 0)$, and that the implementation rule specifies $(p_1, p_2, p_3) = (\frac{2}{3}, \frac{1}{3}, 0)$, where $p_k = \mathsf{P}\{w|r_w = k\}$. If the DM reports their true preferences i.e., $(r_x, r_y, r_z) = (1, 2, 3)$, they face lottery $l = (\frac{2}{3}, x; \frac{1}{3}, y)$, with corresponding utility $U(l) = E_l[v] + \alpha Var_l[v] = 350 + \alpha 20000$; if they instead report $(r_x, r_y, r_z) = (1, 3, 2)$, thus facing lottery $l' = (\frac{2}{3}, x; \frac{1}{3}, z)$, $U(l') = E_{l'}[v] + \alpha Var_{l'}[v] = 300 + \alpha 45000$. Thus, $U(l) > U(l') \iff \alpha < \frac{1}{500}$.

Incentive compatibility of Step 2 The BDM mechanism is not incentive compatible if the DM directly values the uncertainty they face. Depending on the DM's true valuation v, the bias in the report $\tilde{v}$ will be either positive or negative. Below I show this point for $\Psi(\mathbf{p}, \mathbf{v}) = \sum_{\omega \in \Omega} p_\omega(1 - p_\omega)$, but the intuition holds more generally. Let $V \sim \mathcal{U}_{[0,500]}$ be the random compensation. There are two states of the world: $\omega_1 =$ "V is less than $\tilde{v}$" (the DM gets a trip worth v to them) and $\omega_2 =$ "V is greater than or equal to $\tilde{v}$" (the DM gets £V). The DM solves

$$\max_{\tilde{v} \in [0,500]} \mathsf{P}\{V < \tilde{v}\}v + [1 - \mathsf{P}\{V < \tilde{v}\}]E[V|V \geq \tilde{v}] + 2\alpha \mathsf{P}\{V < \tilde{v}\}[1 - \mathsf{P}\{V < \tilde{v}\}]$$

$$\iff \max_{\tilde{v} \in [0,500]} \frac{\tilde{v}}{500}v + \left(1 - \frac{\tilde{v}}{500}\right) \frac{\int_{\tilde{v}}^{500} \frac{V}{500} dV}{\left(1 - \frac{\tilde{v}}{500}\right)} + 2\alpha \frac{\tilde{v}}{500}\left(1 - \frac{\tilde{v}}{500}\right)$$

The FOC is:

$$\frac{v}{500} - \frac{\tilde{v}}{500} + \frac{2\alpha}{500} - \frac{4\alpha\tilde{v}}{500^2} = 0$$

$$\boxed{\tilde{v} = p(\alpha)v + [1 - p(\alpha)]250 \text{ where } p(\alpha) := \tfrac{1}{1+\alpha/125}}$$

It follows immediately that $\tilde{v} > v \iff v < 250$. In other words, the bias in reporting is positive (negative) if the true valuation v is lower (higher) than the midpoint of the support of the distribution of V. Furthermore, $\lim_{\alpha \to 0} \tilde{v} = v$ and $\lim_{\alpha \to \infty} \tilde{v} = 250$, which is intuitive since reporting £250 gives a 50/50 chance of receiving either the compensation or the trip and thus maximizes the uncertainty about the prize.

Although closed-form solutions do not necessarily exist for other measures of uncertainty, I give sufficient conditions on Ψ for the optimal $\tilde{v}$ to be monotone in v. Let $F(\tilde{v}) := \mathsf{P}\{V \leq \tilde{v}\}$ where F is the CDF of the uniform on $[0, 500]$ and consider a DM who chooses $\tilde{v}$ to maximize $U(\tilde{v}; v) = F(\tilde{v})v + [1 - F(\tilde{v})]E[V|V > \tilde{v}] + \alpha\Psi(\tilde{v}, v)$ for some differentiable function Ψ.[10] Using Leibniz rule, the FOC w.r.t. $\tilde{v}$ is:

$$f(\tilde{v})v - f(\tilde{v})\tilde{v} + \alpha \frac{\partial \Psi}{\partial \tilde{v}} = 0$$

Differentiating the FOC w.r.t. v, treating $\tilde{v}$ as an implicit function of v, and remem-

[10] Since V is a continuous random variable, $\mathsf{P}\{V \leq \tilde{v}\} = \mathsf{P}\{V < \tilde{v}\}$. Thus, the problem is equivalent regardless of how the cutoff $V = \tilde{v}$ is treated.

bering $f(\tilde{v}) = \frac{1}{500}$ yields:

$$f(\tilde{v}) - f(\tilde{v})\frac{\partial \tilde{v}}{\partial v} + \alpha \frac{\partial^2 \Psi}{\partial \tilde{v}^2}\frac{\partial \tilde{v}}{\partial v} + \alpha \frac{\partial^2 \Psi}{\partial \tilde{v} \partial v} = 0$$

$$\iff \frac{\partial \tilde{v}}{\partial v} = \frac{f(\tilde{v}) + \alpha \frac{\partial^2 \Psi}{\partial \tilde{v} \partial v}}{f(\tilde{v}) - \alpha \frac{\partial^2 \Psi}{\partial \tilde{v}^2}}$$

Assuming $\alpha > 0$, a sufficient condition for $\frac{\partial \tilde{v}}{\partial v} > 0$ is (i) $\frac{\partial^2 \Psi}{\partial \tilde{v} \partial v} \geq 0$ and (ii) $\frac{\partial^2 \Psi}{\partial \tilde{v}^2} \leq 0$. Part (ii) is satisfied for all measures of uncertainty (concavity in $\mathbf{p}$), while (i) is satisfied for all measures of uncertainty that do not depend on the distribution of valuations (e.g., entropy, residual variance, support); however, this condition is violated for measures such as the variance in valuations.

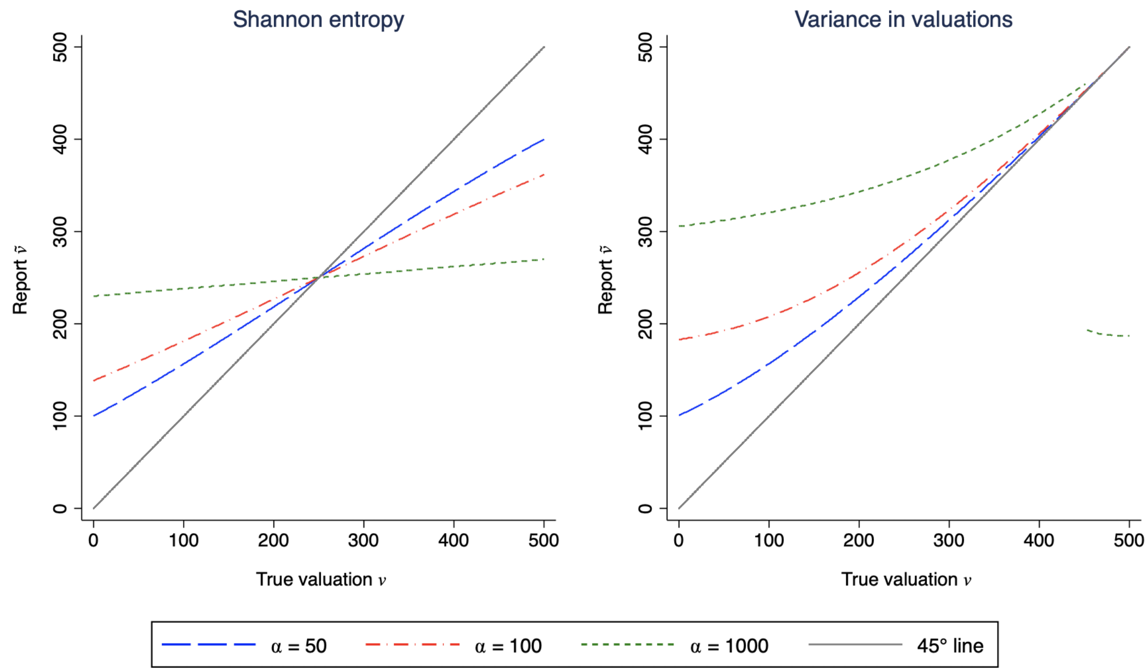
Figure B1 presents the optimal report $\tilde{v}$ as a function of v when the measure of uncertainty is (i) the Shannon entropy (left panel); (ii) the variance in valuations (right panel).[11] To facilitate comparisons, each measure is normalized by its maximum value. The optimal report for the entropy converges to 250 regardless of v as α becomes very large. The optimal report is fairly different when the DM cares instead about the variance in valuations. The bias is in general positive and converges to 0 as v increases provided α is not too large; monotonicity in v is however not guaranteed: as α gets large, the optimal report is in fact a discontinuous function of v.

In summary, the BDM mechanism recovers a biased estimate of the DM's true valuations. The bias is positive for low valuations and negative for high valuations. As a result, the distribution of reports is more compressed than the true distribution. Nevertheless, for most measures of uncertainty considered in this paper, $\tilde{v}$ is a strictly increasing function of v. In other words, the ranking implied by the distribution of reports coincides with the ranking that uses the true valuations, and should also coincide with the ordinal ranking provided in Step 1.

Practical implications and empirical tests If respondents used the elicitation procedure as an additional randomization device, the findings of this paper will understate the importance of preferences for randomization. In addition, given the compression in the reports induced by the BDM mechanism, the observed SD violations will appear less dramatic than they truly are (leading to underestimating $|\alpha|$). How-

[11] The residual variance case is almost identical to the Shannon entropy case.

Figure B1: Optimal report $\tilde{v}$



Notes: Optimal report $\tilde{v}$ as a function of $v \in [0, 500]$ assuming $U(\tilde{v}; v) = \left(\frac{\tilde{v}}{500}\right) v + \left(1 - \frac{\tilde{v}}{500}\right)\left(\frac{\tilde{v}+500}{2}\right) + \alpha\Psi(\tilde{v}, v)$, where $\Psi(\tilde{v}, v) = H(\tilde{v}, v) / \max_{\tilde{v},v} H(\tilde{v}, v)$ and $\alpha \in \{50, 100, 1000\}$. In the left panel, $H(\tilde{v}, v) = -\left(\frac{\tilde{v}}{500}\right) \ln\left(\frac{\tilde{v}}{500}\right) - \left(1 - \frac{\tilde{v}}{500}\right) \ln\left(1 - \frac{\tilde{v}}{500}\right)$ (Shannon entropy); in the right panel, $H(\tilde{v}, v) = [\left(\frac{\tilde{v}}{500}\right)\left(1 - \frac{\tilde{v}}{500}\right)\left(v - \frac{\tilde{v}+500}{2}\right)^2]$ (variance in valuations).

ever, the uncertainty induced by the elicitation mechanism was short-lived (resolved at the end of the survey), thus limiting the anticipatory utility benefits from randomization. Nevertheless, I performed an additional empirical test: if respondents with a preference for randomization used the BDM mechanism as a randomization device, the distribution of reported valuations should be more compressed for respondents with a higher propensity to violate stochastic dominance in favor of randomization; in other words, there should be a negative relationship between the prevalence of SD violations on $\mathcal{D}^+$ or $\mathcal{B}^+$ and the standard deviation of the reported valuations, std($\{\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_{10}\}$). Instead, I find that the correlation is insignificant and, if anything, positive. There is also no relationship between preference for randomization in the design task and compression in the reported valuations.

B.2 Data on ranking and valuation of the destinations

Figure B2: Distribution of the ranks 1-10 assigned to each destination

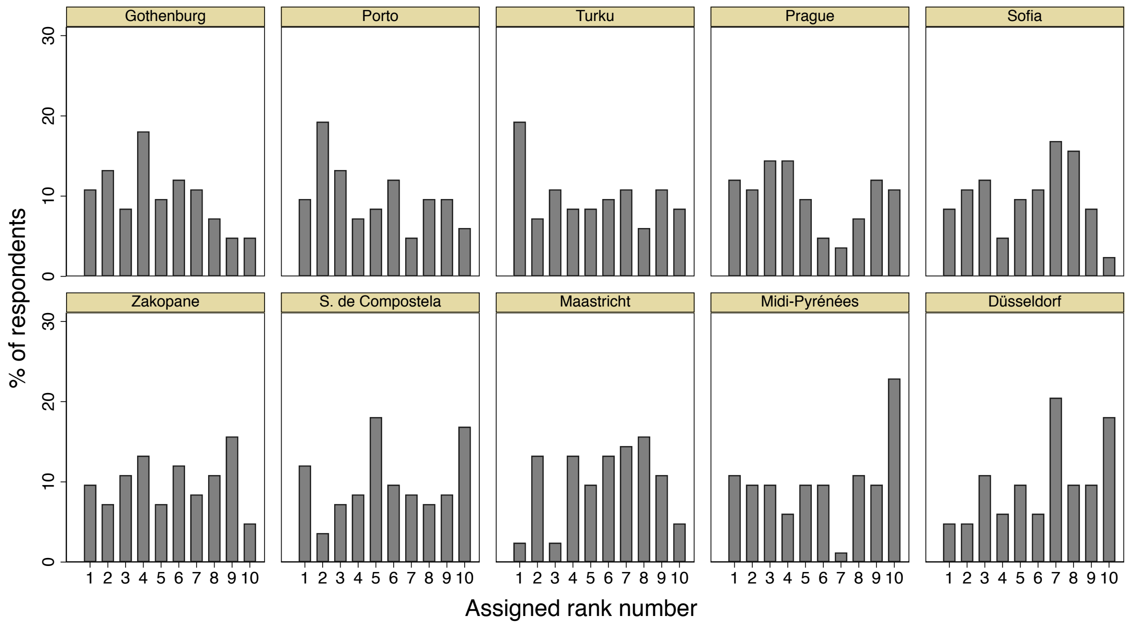
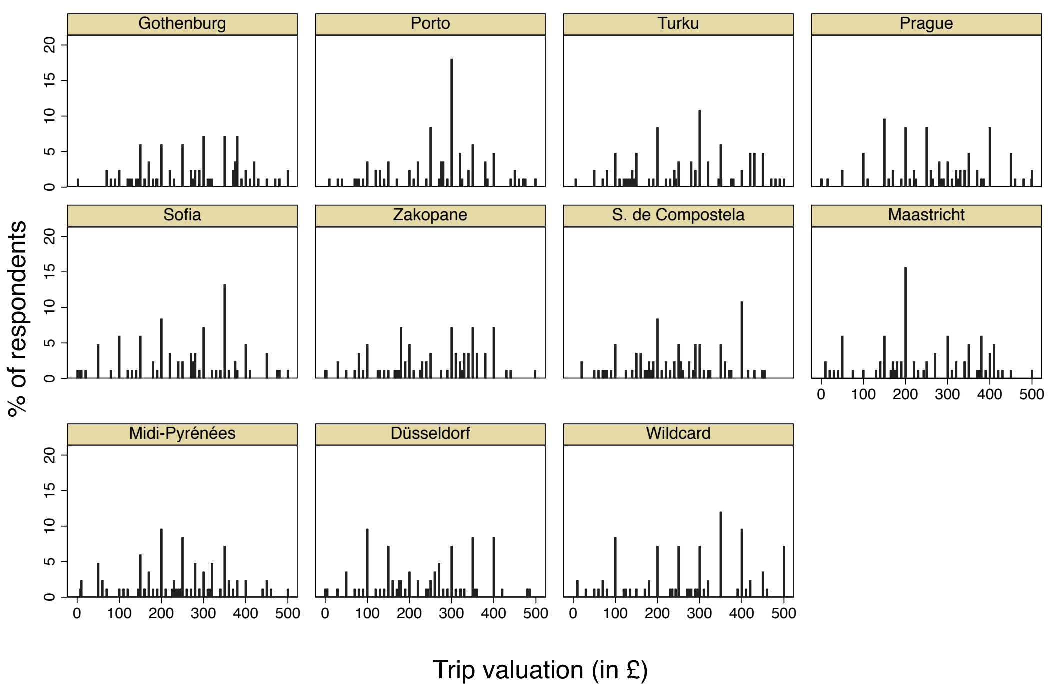
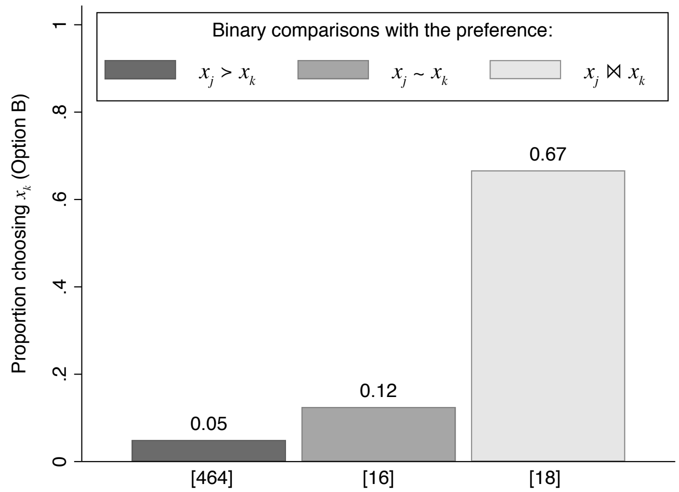


Figure B3: Distribution of respondents' valuations by destination ($v \in [0, 500]$)



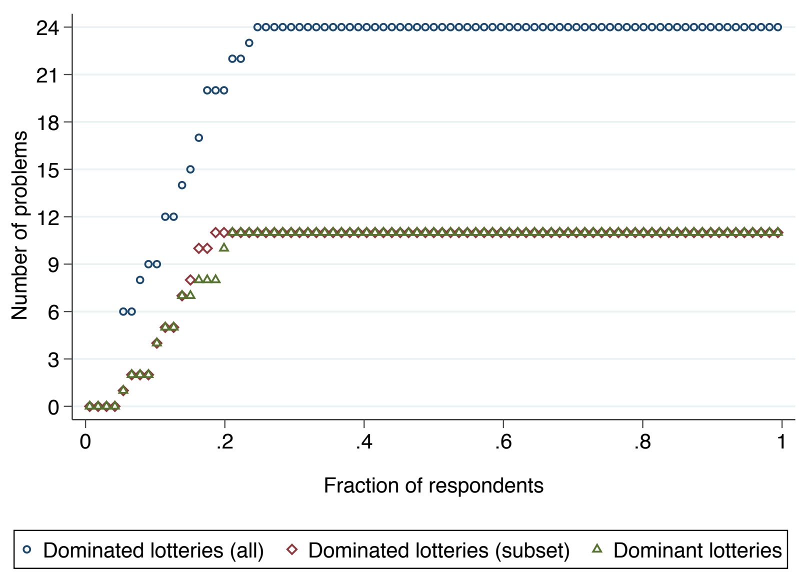
B.3 Preference consistency and prevalence of SD violations

Figure B4: Proportion choosing x_k from $\{x_j, x_k\}$ depending on elicited preference $\succeq$



Notes: Number in square brackets under each bar equal to number of respondents × number of binary comparisons at which a respondent had a strict preference $x_j \succ x_k$, an indifference $x_j \sim x_k$, or an incomparability $x_j \bowtie x_k$. The total number of observations is 498 (= 83 × 6). For instance, the option x_k was chosen from $\{x_j, x_k\}$ in 5% of the 464 cases in which $x_j \succ x_k$ on any comparison.

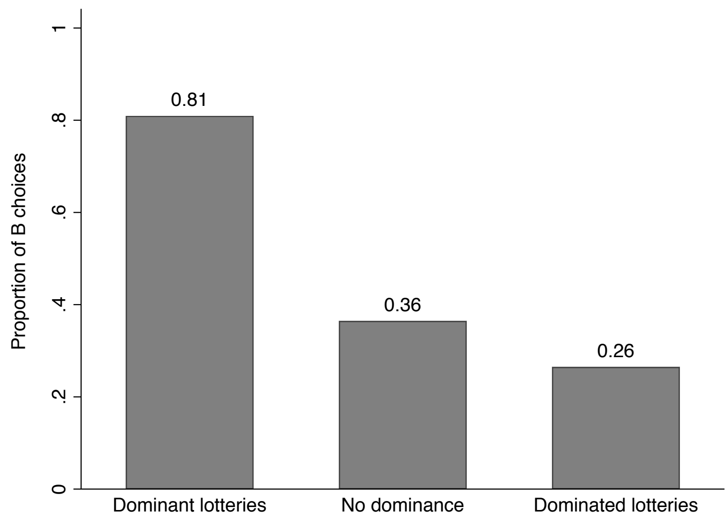
Figure B5: Number of clear dominance problems per respondent



Notes: Number of clear dominance problems calculated for each set ($\mathcal{D}^+$, $\mathcal{B}^+$ and $\mathcal{B}^-$) by using the unambiguous relation $\succeq^*$. The maximum number is 24 for $\mathcal{D}^+$ and 11 for $\mathcal{B}^+$ and $\mathcal{B}^-$.

B.4 More information on the size and shape of SD violations

Figure B6: Frequency of B choices (lottery) by type of problem



Notes: Proportions calculated from the total number of binary comparisons belonging to each class of problems. For instance, the denominator for the fraction of B choices in the set of 11 problems with dominant lotteries is 11×83 respondents $= 913$.

Figure B7: Choices between x_5 and $(x_1, p; x_{10}, 1-p)$

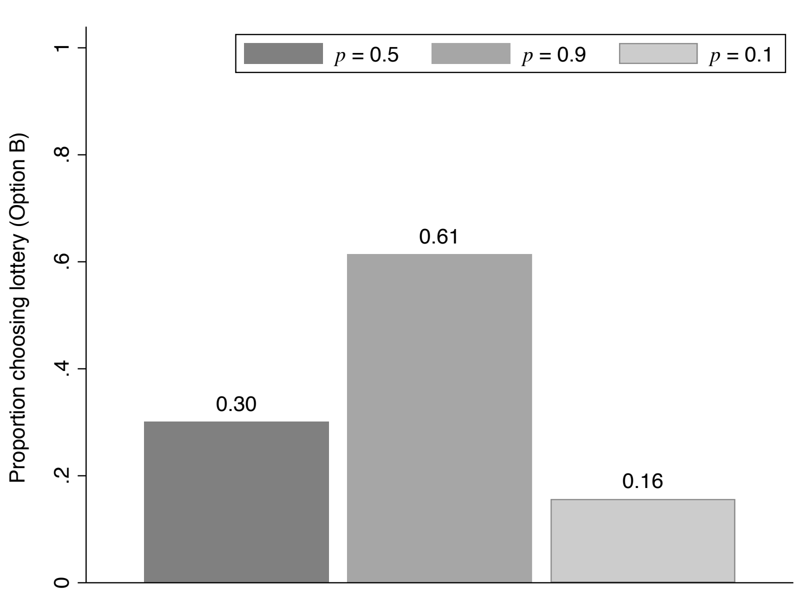


Table B1: Expected value difference between Option A and B

	Mean	SD	p25	p50	p75	Min	Max	N
SD violations on $\mathcal{D}^+$ (dominated lotteries)								
x_1 vs. $(x_1, 0.5; x_2, 0.5)$	18.4	13.3	10	15	25	2	55	39
x_1 vs. $(x_2, 0.5; x_3, 0.5)$	27.2	12.9	18	25	30	10	55	26
x_1 vs. $(x_1, 0.5; x_2, 0.5)$	52.5	31.2	30	41	75	15	125	20
x_1 vs. $(x_1, 0.5; x_5, 0.5)$	76.3	27.1	50	75	100	40	125	13
x_1 vs. $(x_1, 0.5; x_{10}, 0.5)$	93.0	42.4	50	100	120	50	145	5
x_1 vs. $(x_1, 0.9; x_2, 0.1)$	4.2	3.0	2	3	5	0	11	26
x_1 vs. $(x_1, 0.9; x_3, 0.1)$	7.2	3.5	5	6	10	3	15	21
x_1 vs. $(x_2, 0.9; x_3, 0.1)$	35.5	24.8	21	27	42	8	105	14
x_1 vs. $(x_1, 0.9; x_5, 0.1)$	13.4	6.2	10	11	17	5	25	11
x_1 vs. $(x_1, 0.9; x_{10}, 0.1)$	20.2	8.5	10	22	27	10	30	6
x_1 vs. $(x_1, 0.1; x_2, 0.9)$	36.0	26.6	18	27	45	9	90	13
x_1 vs. $(x_1, 0.1; x_3, 0.9)$	72.0	43.0	40	54	112	27	135	8
x_1 vs. $(x_2, 0.1; x_3, 0.9)$	71.6	44.5	38	55	95	29	145	9
x_1 vs. $(x_1, 0.1; x_5, 0.9)$	134.4	62.8	90	148	189	45	225	7
x_1 vs. $(x_1, 0.1; x_{10}, 0.9)$	243.0	38.2	216	243	270	216	270	2
x_1 vs. $(x_1, 0.4; x_2, 0.4; x_3, 0.3)$	30.1	21.1	18	22	36	6	98	26
x_1 vs. $(x_1, 0.4; x_3, 0.4; x_5, 0.3)$	52.6	26.4	33	45	69	15	108	14
x_1 vs. $(x_2, 0.4; x_3, 0.4; x_4, 0.3)$	54.0	34.7	34	46	65	12	145	19
x_5 vs. $(x_6, 0.4; x_7, 0.4; x_8, 0.3)$	38.4	20.1	19	36	55	3	71	22
x_1 vs. $(x_1, 0.2; x_2, 0.2; x_3, 0.2; x_4, 0.2; x_5, 0.2)$	59.4	34.3	38	49	72	12	161	26
x_1 vs. $(x_1, 0.2; x_3, 0.2; x_5, 0.2; x_7, 0.2; x_{10}, 0.2)$	99.7	39.3	78	100	131	28	147	7
x_1 vs. $(x_6, 0.2; x_7, 0.2; x_8, 0.2; x_9, 0.2; x_{10}, 0.2)$	170.0	75.8	136	160	214	44	282	7
x_5 vs. $(x_6, 0.2; x_7, 0.2; x_8, 0.2; x_9, 0.2; x_{10}, 0.2)$	58.2	25.5	40	60	76	10	110	15
x_1 vs. $(x_1, 0.1; x_2, 0.1; ...; x_9, 0.1; x_{10}, 0.1)$	94.6	41.7	73	96	134	28	140	6
SD violations on $\mathcal{D}^-$ (dominant lotteries)								
x_2 vs. $(x_1, 0.5; x_2, 0.5)$	18.5	11.9	10	15	25	5	50	15
x_3 vs. $(x_1, 0.5; x_3, 0.5)$	36.2	29.1	10	28	50	10	100	10
x_3 vs. $(x_1, 0.5; x_2, 0.5)$	48.5	41.5	20	35	70	15	150	10
x_5 vs. $(x_1, 0.5; x_5, 0.5)$	52.5	33.5	25	39	88	25	100	6
x_{10} vs. $(x_1, 0.5; x_{10}, 0.5)$	103.2	97.2	40	55	215	40	215	3
x_2 vs. $(x_1, 0.9; x_2, 0.1)$	37.2	23.3	18	36	45	9	90	18
x_3 vs. $(x_1, 0.9; x_3, 0.1)$	56.2	34.2	27	45	90	18	112	10
x_3 vs. $(x_1, 0.9; x_2, 0.1)$	55.4	34.9	19	38	95	19	95	7
x_2 vs. $(x_1, 0.1; x_2, 0.9)$	3.8	2.0	2	4	5	1	10	23
x_3 vs. $(x_1, 0.1; x_3, 0.9)$	7.0	4.0	5	6	10	2	20	22
x_3 vs. $(x_1, 0.1; x_2, 0.9)$	29.3	31.1	11	22	45	5	110	11

Notes: Summary statistics of the absolute difference in monetary value between Option A (sure option) and Option B (lottery) i.e., $|v(x_A) - E_{\mathbf{p}}(v_B)|$ among respondents who violated stochastic dominance according to $\succeq^*$ in the relevant decision problem. The quartiles, minimum and maximum were rounded to the nearest integer amount.

C Decisions in the design task

C.1 Description of lottery-building exercise

After the 45 DPs, respondents were offered to build their own lottery or select a sure destination instead. Those who chose to build a lottery first selected the support and then allocated 100 lottery tickets to the destinations they selected. After that, they were offered to further personalize their lottery by selecting the date at which to reveal the destination from a dropdown menu.[12]

Those who initially preferred a sure destination were asked whether they would now build a lottery if they could choose the date at which to learn the destination. This splits respondents into 3 groups depending on whether they preferred to build a lottery right away, when the delay was fixed (henceforth L1 for "Lottery 1st time"), only after the delay was customizable (henceforth L2 for "Lottery 2nd time"), or not at all, instead preferring a sure destination (henceforth NL for "No Lottery").

Respondents who built a lottery (L1 and L2) and chose a delay (all but "*Today, after I completed the survey*") were then asked how much they value the possibility of keeping the surprise until their chosen date. Valuations were elicited using a Multiple Price List (MPL) with respondents making choices in 7 rows:

Revealing Later	○	○	Revealing Now + £5
Revealing Later	○	○	Revealing Now + £10
Revealing Later	○	○	Revealing Now + £15
Revealing Later	○	○	Revealing Now + £20
Revealing Later	○	○	Revealing Now + £30
Revealing Later	○	○	Revealing Now + £40
Revealing Later	○	○	Revealing Now + £50

where "Revealing Later" is the respondent's chosen date and "Revealing Now" means that the destination is revealed at the end of the survey. I proxy a respondent's WTP for delay by using the last row at which s/he selected "Revealing Later", a lower bound on the respondent's true valuation.

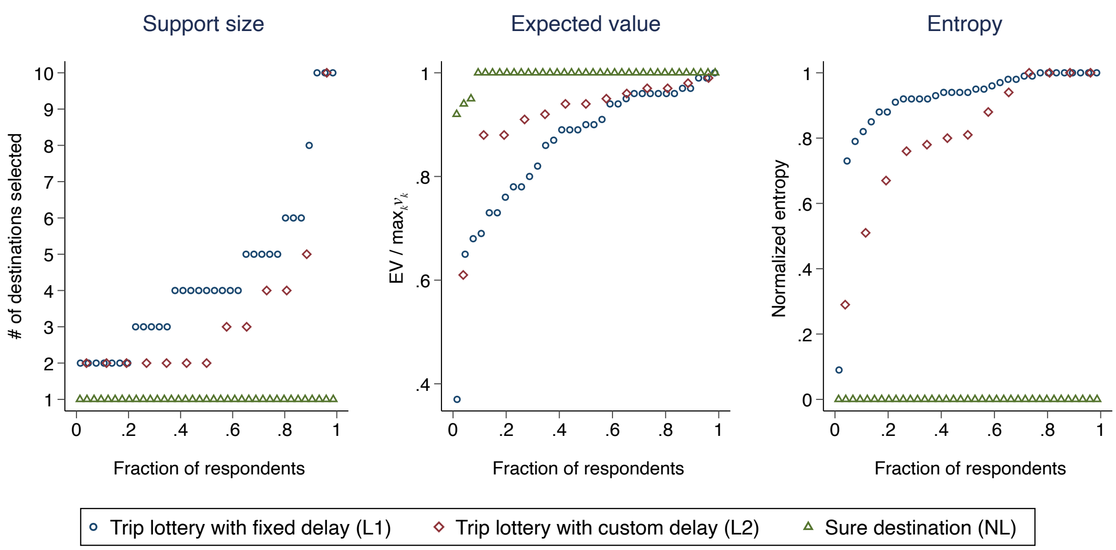
[12]The options were: "*Today, after I completed the survey*", "*About 48 hours from now*", "*About one week from now*", "*About 2 weeks from now*", "*About 4 weeks from now*", "*About 4 weeks before going to the airport*", "*About 2 weeks before going to the airport*", "*About 1 week before going to the airport*", "*About 24 hours before going to the airport*", "*On a specific day of my choice*" [specified right after].

C.2 Characteristics of the favorite option

C.2.1 Choice of destination(s) and probability weights

Figure C1 presents the characteristics of the option chosen by each of the 83 respondents. In total, 55% of respondents chose a lottery as their favorite option, 40% (33/83) on the first occasion and 15% (13/83) on the second occasion, with the remaining selecting a sure destination (37/83). Generally speaking, L1 respondents selected lotteries that have more uncertainty than L2 respondents and were willing to sacrifice more in terms of expected value from the trip. Reassuringly, over 90% of the 37 respondents who preferred a sure destination selected the option they assigned the highest valuation to; furthermore, at most 8% of total value was sacrificed by the 3 respondents who did not.[13] Among the 46 respondents who chose to build a lottery, 43 selected their rank #1 destination as an element of the support of their lottery.

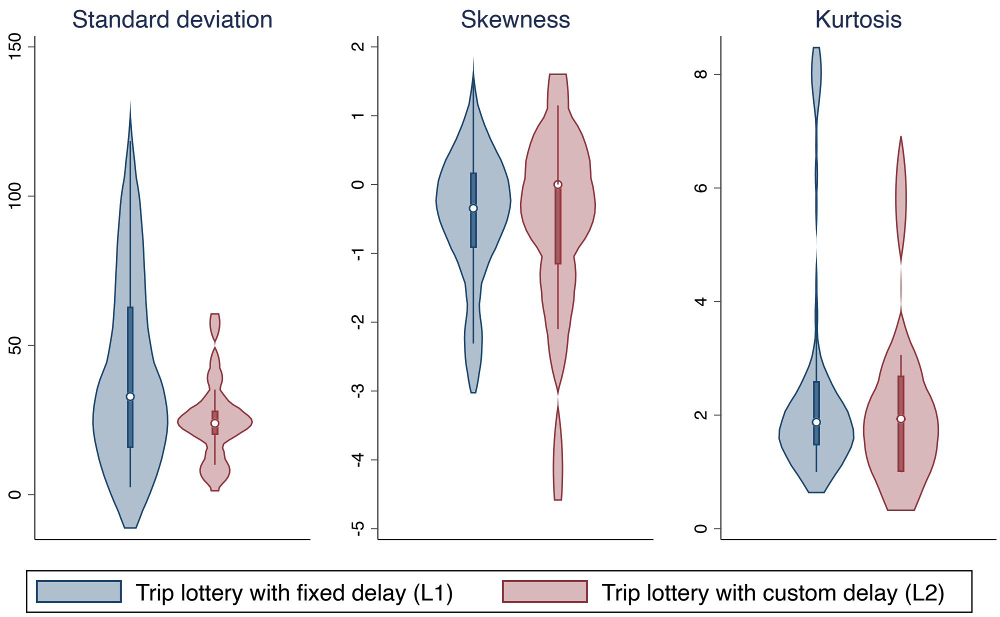
Figure C1: Characteristics of respondents' favorite option



Notes: The 3 panels present quantile plots (using uniform scaling) of (i) the size of the support of $\mathbf{p}$ chosen by respondents; (ii) the expected value of the chosen option relative to the highest valuation, $\mathbf{p} \cdot \mathbf{v} / \max_k v_k$; (iii) the entropy for the chosen probability $\mathbf{p}$ normalized by $\ln(|\text{supp}(\mathbf{p})|)$ so the measure lies between 0 and 1. N = 83.

[13] In total, 84% (31/37) of the respondents who preferred a sure destination selected the one they ranked #1 in Section 1. In 3 of the 6 cases where they did not, they in fact selected the option they valued the most (meaning that the ranking and valuations disagreed).

Figure C2: Second, third and fourth moments of chosen lottery



Notes: Violin plots of the distribution of the second moment (standard deviation), third moment (skew) and fourth moment (kurtosis) of the distribution of valuations for the trip lottery chosen by respondents who selected one as their favorite option. In each plot, the white dot corresponds to the median, the box to the interquartile range, and the spikes extend to the upper- and lower-adjacent values. N = 46 except for one outlier removed for skewness ($\mu_{3,i} = -9.8$) and two outliers were removed for kurtosis ($\mu_{4,i} = 97.6$ and $\mu_{4,i} = 18.1$), in order to improve visibility.

Figure C2 presents the 2nd, 3rd and 4th moments of the distribution of valuations for the chosen lottery (separately for the L1 and L2 groups). Lotteries tend to be negatively skewed, putting higher weight on options ranked/valued higher. The mean level of skew for the two groups combined is $\bar{\mu}_3 = -0.75$, significantly different from 0 (p = 0.006), with 59% of lotteries being negatively skewed, 28% being positively skewed, and 13% being symmetric (i.e., $\mu_{3,i} = 0$). This finding is consistent with a $U_{\alpha,\Psi}$-representation in which the measure of uncertainty Ψ depends only on $\mathbf{p}$, such as the Shannon entropy or the residual variance. In particular, in both models, the following monotonicity property should hold:

$$\text{For any } x_k, x_{k+1} \in X, \; p_k \geq p_{k+1}$$

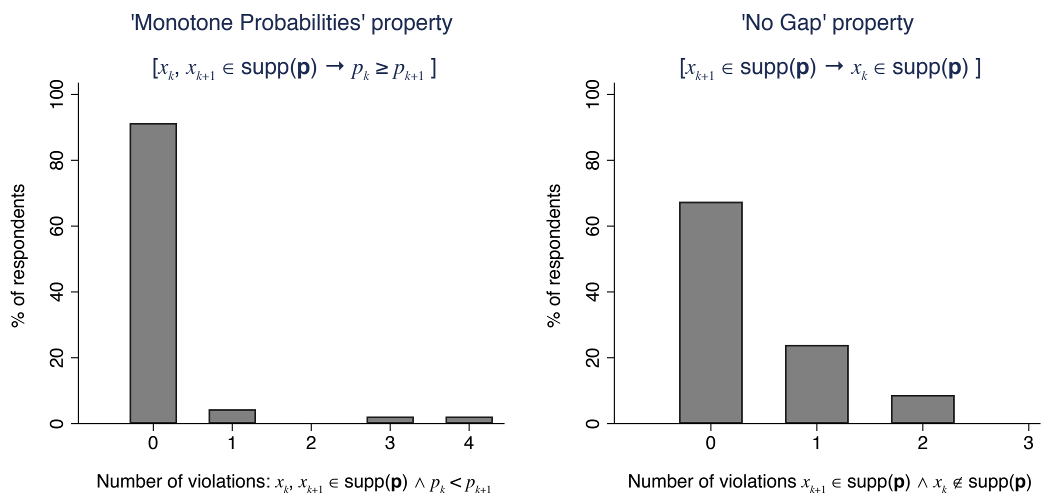
(where $k \in \{1, 2, ..., 10\}$ corresponds to the rank number in the respondent's ranking). Among the 46 respondents who built a lottery, 59% systematically assigned a higher probability weight to options they ranked higher.[14] In addition, when deviations from this property occur, they tend to be minimal, with respondents (i) sometimes failing to select one higher-ranked option but assigning monotone weights otherwise and/or (ii) always selecting higher-ranked options but sometimes failing to give them a (weakly) higher weight. To examine these deviations further, I decompose the above monotonicity property into two parts corresponding to cases (i) and (ii):

Monotone Probabilities For any $x_k, x_{k+1} \in \text{supp}(\mathbf{p})$, $p_k \geq p_{k+1}$.

No Gap For any $x_k, x_{k+1} \in X$, if $x_{k+1} \in \text{supp}(\mathbf{p})$ then $x_k \in \text{supp}(\mathbf{p})$.

Figure C3 shows that 'No Gap' is violated more often than 'Monotone Probabilities', although both are infrequent. Again, these findings are more consistent with a $U_{\alpha,\Psi}$-representation where Ψ is independent of $\mathbf{v}$ e.g., a DM who enjoys variance in the trip valuations would violate 'No Gap'.

Figure C3: Violations of Monotone Probabilities and No Gap



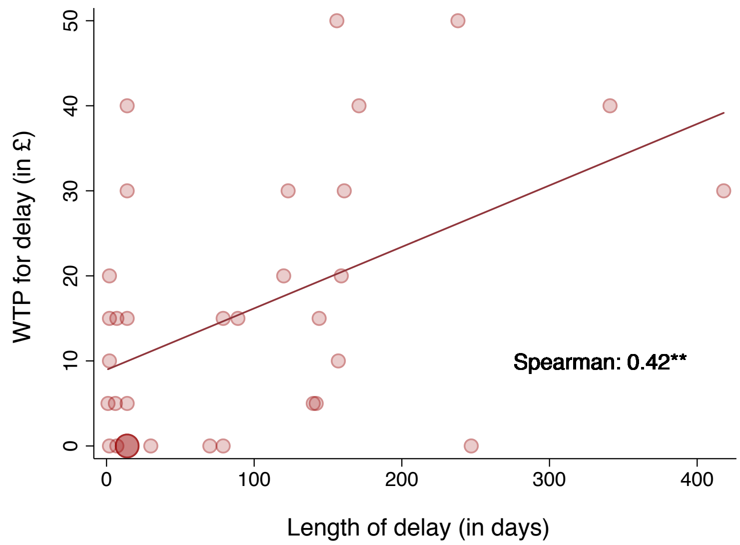
Notes: Distribution of the number of times a designed lottery failed to satisfy the "Monotone Probabilities" and "No Gap" properties. N = 46 (i.e., all respondents who built a lottery).

[14]Note that this is a stronger requirement than requiring that for any $x_k, x_{k+1} \in X$ such that $x_k \succeq x_{k+1}$, then $p_k \geq p_{k+1}$, because valuations might disagree with the ranking (i.e., $v_k < v_{k+1}$) so that $x_k \bowtie x_{k+1}$. However, conditioning only on the set of 40 respondents whose ranking and valuations fully agree, this number only rises to 63% (25/40) and remains nearly identical at 62% (21/34) if one further removes respondents who made direct choices in the binary decision problems that contradicted their ranking (thus focusing on $\succeq^*$).

C.2.2 Preferences for delaying the resolution of uncertainty

Link between delay length and WTP for delay Figure C4 shows a positive and statistically significant relationship between the amount of delay chosen by those who preferred to postpone the revelation of uncertainty (N = 34) and their willingness to pay for delay, with a rank-order correlation of 0.42 (p = 0.013).

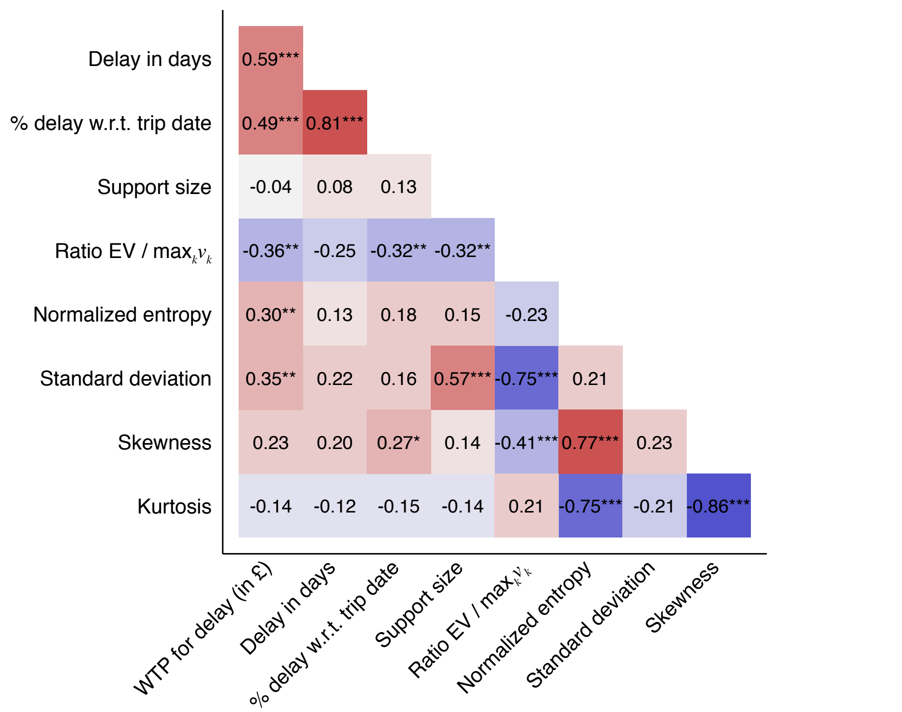
Figure C4: Relationship between WTP for delay and delay length



Notes: ** $p < 0.05$. The red line is from a linear regression between the two variables ($\hat{\beta} = -0.07$, s.e.= 0.02, p = 0.004). N= 34 (based on the respondents' selections in the MPL exercise).

Delay preferences and scope for surprise Figure C5 shows a matrix of Pearson correlation coefficients for the various lottery characteristics chosen by the 46 respondents who build a lottery. Overall, there is some suggestive (non-robust) evidence of a positive relationship between preferences for preserving the surprise of the destination and scope for surprise as measured by the entropy or support size of the lottery and the upside (downside) of getting a higher- (lower-)ranked outcome.

Figure C5: Correlations between lottery characteristics

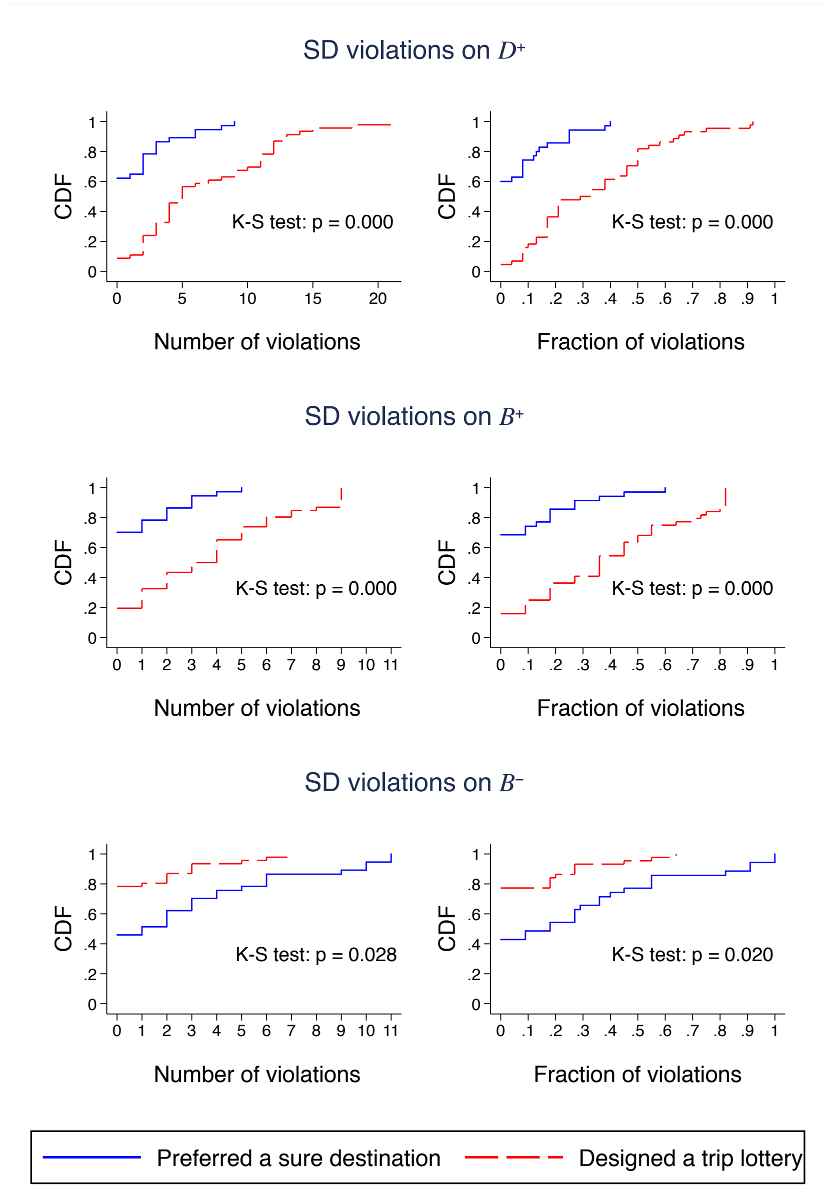


Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The coefficients are Pearson correlations. N = 46.

C.3 Preferences in the binary choice exercise vs. design task

C.3.1 Link between SD violations and choice to design a lottery

Figure C6: Relationship between binary choices and lottery-building exercise

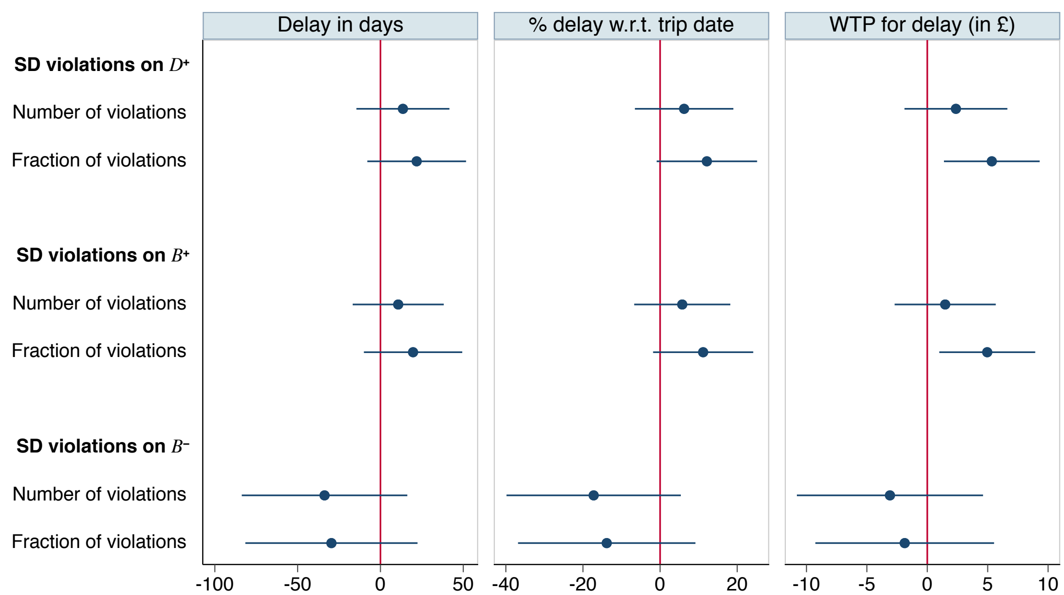


Notes: p-values are from Kolmogorov-Smirnov tests of equality of distributions. N = 37 for the respondent group "Preferred a sure destination" and N = 46 for the "Designed a trip lottery" group. The latter group pools respondents who already preferred to build a trip lottery without customizable delay (i.e., delay fixed at one week before travel; N = 33) and those who preferred to build a lottery once offered to customize the delay (N = 13).

C.3.2 Link between SD violations and preferences for delay

If SD violations on $\mathcal{D}^+$ ($\mathcal{B}^+$) reflect a preference for surprise, they should be more frequent among respondents who preferred discovering their destination later. Similarly, if SD violations on $\mathcal{B}^-$ reflect a preference for certainty, they should be less frequent among those with a preference for delay. Figure C7 presents suggestive evidence in this direction, although the sample is too small to draw conclusions. Correlations have the expected sign i.e., with a stronger preference for delay (as measured by chosen delay in days, % delay relative to trip date, and WTP for delay) being positively (negatively) associated with SD violations on $\mathcal{D}^+$ and $\mathcal{B}^+$ ($\mathcal{B}^-$), but a limited number are even marginally significant. Among them, the Pearson correlation between WTP for delay and the fraction of SD violations on $\mathcal{D}^+$ is 0.39 (p = 0.009).

Figure C7: Correlation between preference for delay and SD violations



Notes: Coefficients are from univariate linear regressions of each measure of preference for delay on the z-scores of the listed variables. Error bars are 95% confidence intervals. For instance, a one standard-deviation increase (relative to the sample mean) in the fraction of SD violations exhibited on $\mathcal{D}^+$ is associated with an increase of 21.9 days (95% CI: [-7.9, 51.7]) in the chosen delay, on average. N = 46 (44) for regressions with number (fraction) of SD violations on $\mathcal{D}^+$, $\mathcal{B}^+$ and $\mathcal{B}^-$.

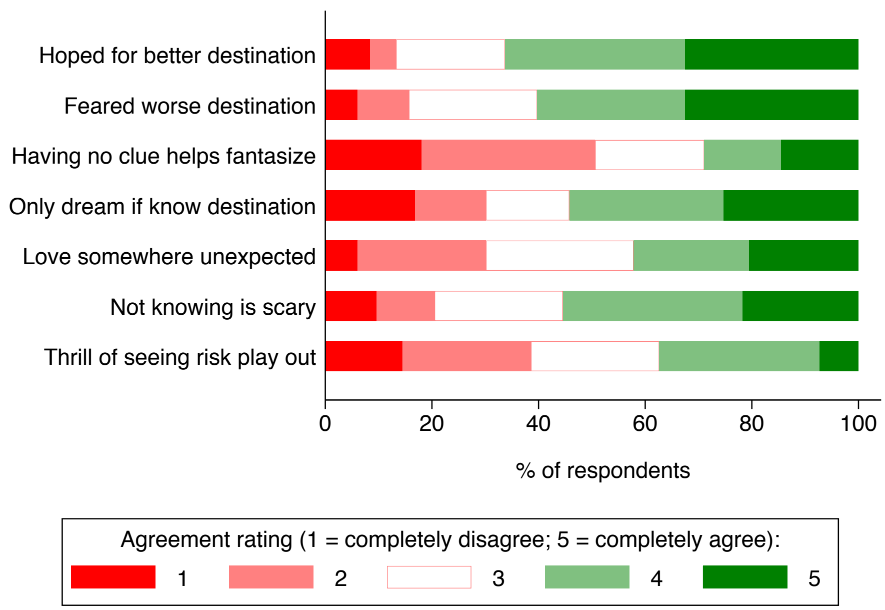
D Drivers of preferences for surprise trips

This section presents additional information on the psychological drivers of respondents' preferences for surprise trips.

D.1 Psychological motives behind wildcard valuations

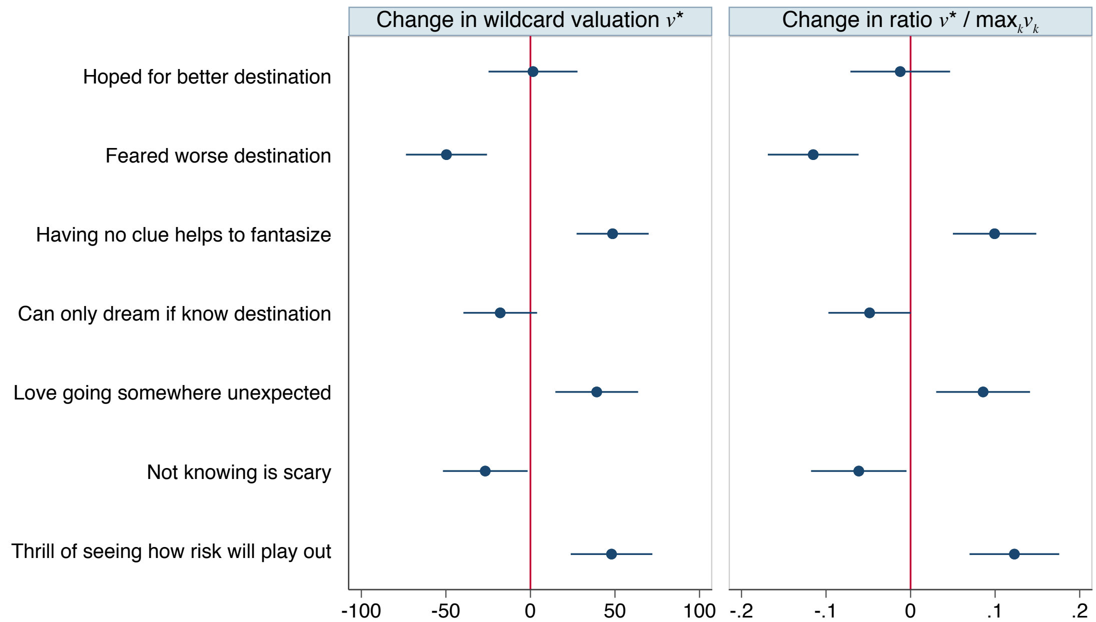
After respondents submitted their valuations for the wildcard trip, they were asked to rate their agreement with various factors. As shown in Figure D1, respondents generally agreed that they hoped for a better destination and feared a worse destination, that dreaming about one's holidays requires to know the destination, and that not knowing the destination feels scary; conversely, most respondents disagreed with the statement that having no clue about the destination helps to fantasize about any type of vacation. Views were more divided on the thrill of waiting to see how risk will play out and on the love of going somewhere they would have never thought about.

Figure D1: Agreement with reasons



Notes: Distribution of answers to the question: "*So we know what factors helped you determine your valuation of the wildcard trip, please rate the extent to which you agree with the following statements:*" (1 = completely disagree; 5 = completely agree). The statements were shortened to make space on the figure; see instructions for the full statements. N = 83.

Figure D2: Determinants of wildcard valuation



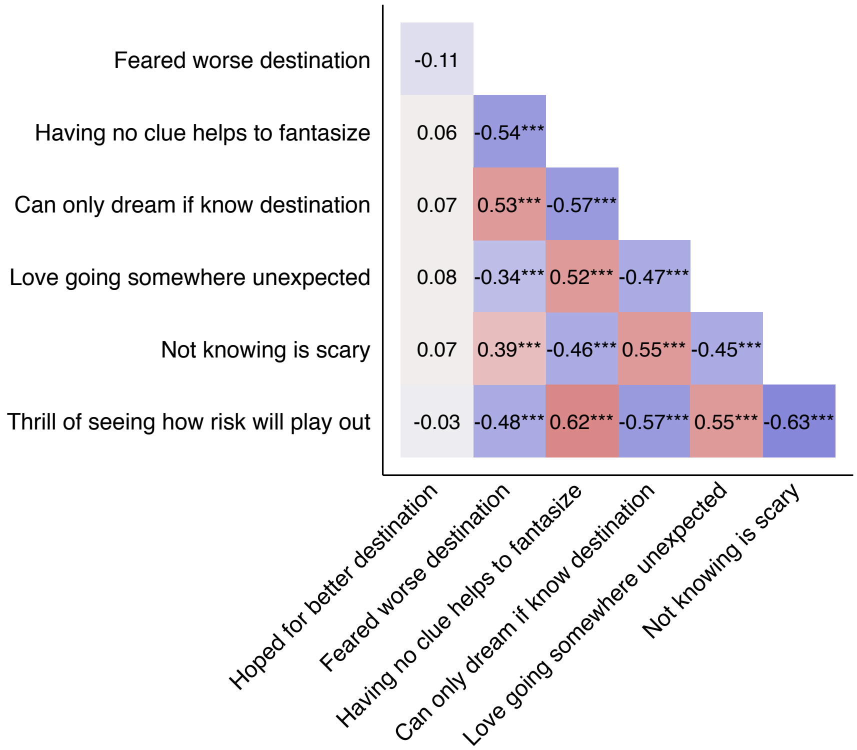
Notes: Left panel: Two-limit Tobit regressions of wildcard trip valuation v^* (in £) on each of 7 ratings (where 1 = "completely disagree" and 5 = "completely agree") entered separately; right panel: linear regression of the ratio $v^*/\max_k v_k$ (where $\max_k v_k$ is the highest trip valuation that the respondent entered for the 10 destinations) on these ratings. N = 83.

While certain reasons were more popular than others, they might not be necessarily the ones that explain the observed variation. In fact, Figure D2 shows that hoping for a better destination has no predictive power, but fear of a worse destination does: a one-point increase in agreement with the statement "*I feared I would get a trip I like less than the other destinations.*" is associated with a £49.6 decrease (95% CI: [-73.6, -25.7]) in valuations on average. In terms of positive factors, the ability to fantasize about different worlds and the thrill of risk appear to be the strongest predictors, with a £48.6 increase (95% CI: [27.3, 69.9]) on average for the former, and a £48.0 increase (95% CI: [23.8, 72.1]) for the latter. These effects correspond to a 10 to 15-percentage points change relative to the maximum valuation (right panel of Figure D2).

Figure D3 shows that these various factors tend to be fairly highly correlated, with most Pearson correlation coefficients in the range [0.4, 0.6], at the exception of "Hope for better destination", which is not correlated with any of the other factors.[15]

[15] As a note, multivariate regressions with all factors entered as predictors produces some counterintuitive findings (positive and significant coefficient on "Can only dream if know destination")

Figure D3: Correlation in attitudes towards wildcard trip



Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The coefficients are Pearson correlations. The statements were shortened for space reasons; see instructions for the full statements. N = 83.

Beliefs about possible outcomes After entering their valuations, respondents were asked to estimate their chances of preferring the wildcard trip to the destinations they ranked first and last, x_1 and x_{10}. Denote those estimates by $\mu_1 := \Pr(x_* \succ x_1)$ and $\mu_{10} := \Pr(x_* \succ x_{10})$. Respondents were overall very optimistic that the wildcard trip would be better than x_{10}, with the majority of respondents assigning at least 80% chance to this possibility. They were also quite optimistic about their chances of obtaining a trip they like better than x_1, with nearly 40% of respondents expecting this to happen with 50% chance or more. There is a strong positive correlation between a respondent's wildcard valuation v^* and their belief μ_1 (Pearson $\rho = 0.50$, $p < 0.001$), but no correlation between v^* and μ_{10}.[16]

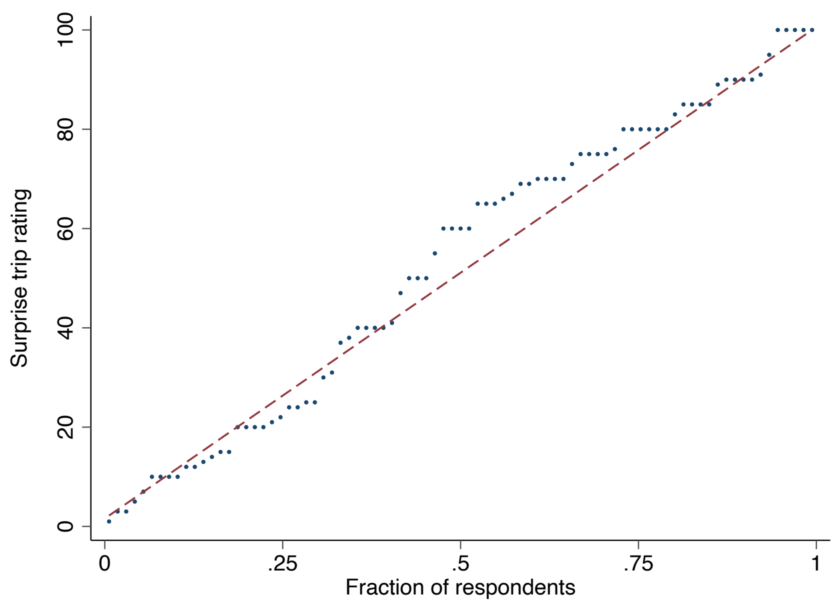
and, at the exception of "Feared worse destination", the estimated effects for the other reasons are non robust and sometimes depend on the choice of outcome variable (v^* or $v^*/\max_k v_k$).

[16]To gain more insights on the role of beliefs, respondents were also asked prior to entering their valuations to estimate the number of destinations in the company's database and, among them, the number they thought they might be able to guess. Subsequently, half of the respondents were told

D.2 Attitudes towards surprise trips

Distribution of ratings At the end of the survey, respondents were asked to rate on a scale from 0 to 100 how much overall they liked the concept of "surprise trip" presented in the study (0 = "Did not like it at all"; 100 = "Absolutely loved it"). In line with the choice data, views tended to be positive but with large heterogeneity: the mean (median) rating was 53.3 (60), with a standard deviation of 30.6. The distribution of the rating is shown in Figure D4.

Figure D4: Distribution of surprise trip ratings (0-100)

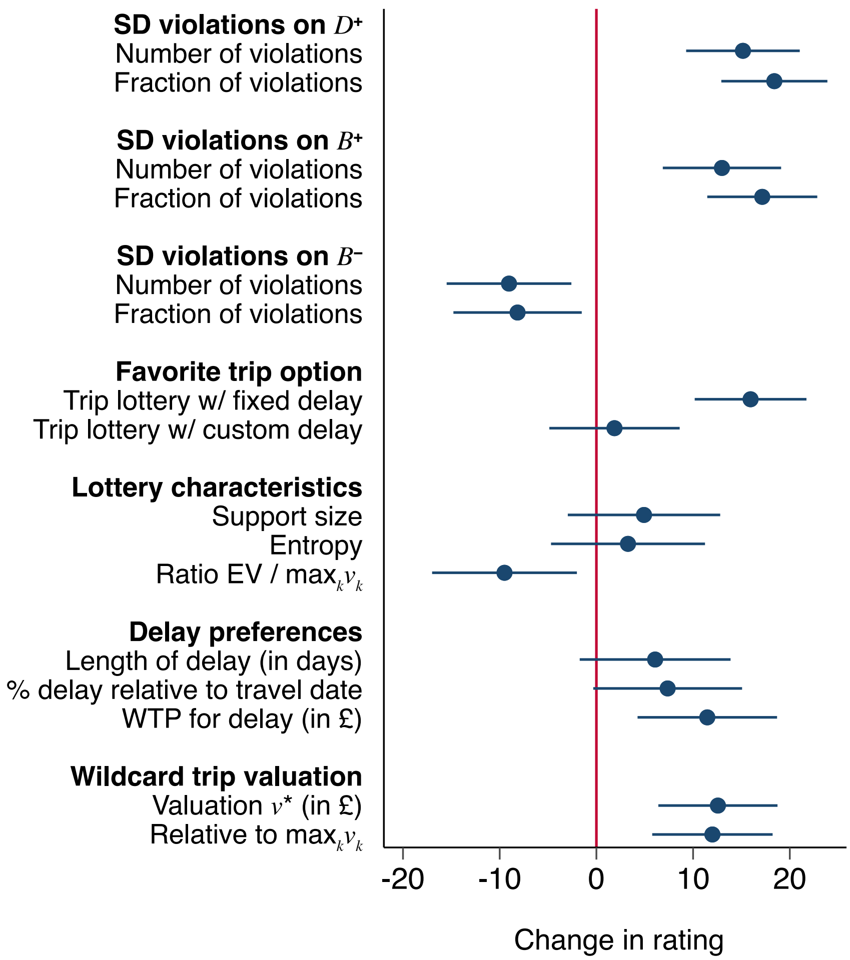


Notes: Quantile plot of the distribution of surprise trip ratings (0-100) against the 45° line. Ratings provided when answering the question "*Overall, how much did you like the concept of "surprise trip" presented in this study?*, with 0 = "Did not like it at all"; 100 = "Absolutely loved it".

Correlates of the rating As expected, this rating correlates very well with the various measures of preference for randomization and delay considered in this paper (Figure D5).

the exact number (> 20,000 destinations); this was done to test whether a larger perceived state space (implying more surprise) translates into higher valuations. Overall, I find no impact of this treatment manipulation.

Figure D5: Correlates of the surprise trip rating (0-100)



Notes: Coefficients are from simple linear regressions of `Surprise trip rating (0-100)` on the z-scores of the listed variables. Error bars are 95% confidence intervals. For instance, a one standard-deviation increase (relative to the sample mean) in the wildcard valuation v^* entered by a respondent is associated with a 12.6 point increase (95% CI: [6.4, 18.7]) in the surprise trip rating. N = 83 (79) for the number (fraction) of SD violations on $\mathcal{D}^+$, $\mathcal{B}^+$ and $\mathcal{B}^-$; N = 83 for the variables under "Favorite trip option" and "Wildcard trip valuation"; N = 46 for the variables pertaining to "Lottery characteristics" and "Delay preferences".

E Knowledge and preferences over trip attributes

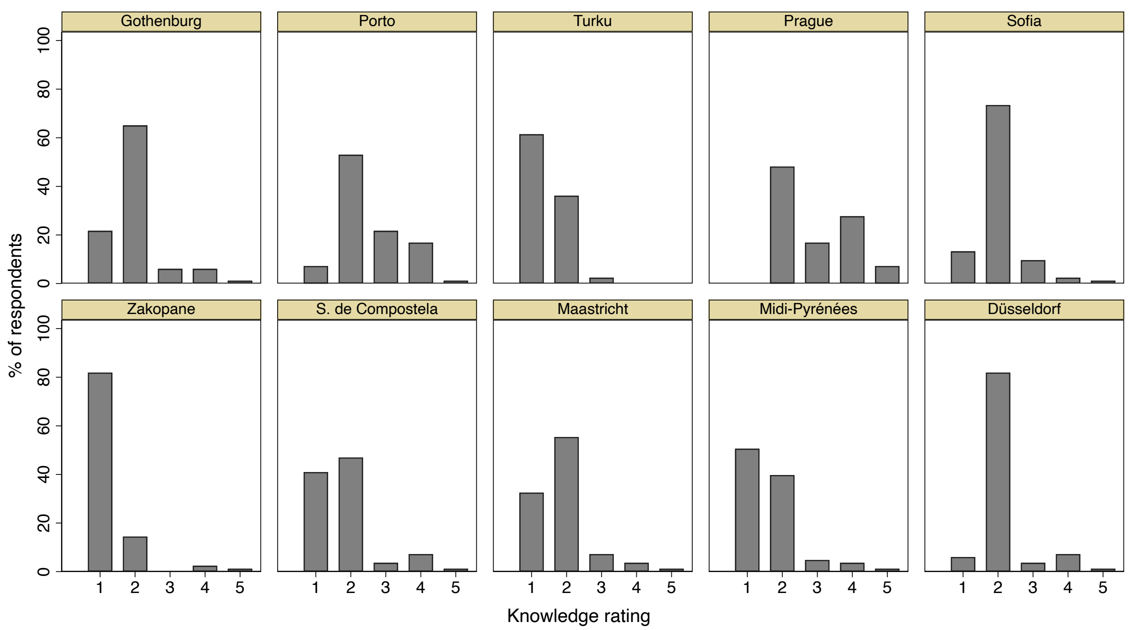
E.1 Knowledge of trip destinations

Right after seeing the description of each trip, participants were asked to rate their knowledge of each destination on a scale from 1 to 5 with:

1. "*Very unfamiliar - Never heard of it*"
2. "*Fairly unfamiliar - Heard of it but never been*"
3. "*Neither familiar nor unfamiliar - Been once but barely remember /saw anything*"
4. "*Fairly familiar - Been a few times and/or visited quite a bit*"
5. "*Very familiar - Been many times and/or visited a lot*"

The most common ratings were 1 or 2 for virtually all destinations, indicating limited knowledge and thus many opportunities for a novel experience.

Figure E1: Distribution of knowledge ratings



E.2 Preferences over trip attributes

At the end of the survey, participants were asked to rate where they locate on the spectrum for various trip attributes such as the weather, the accommodation, etc. Figure E2 shows the distribution of responses for each attribute. In general, participants have a preference for going to renowned but new destinations. Figure E3 shows the link between attribute preferences and valuation of each destination.

Figure E2: Preferences over trip attributes

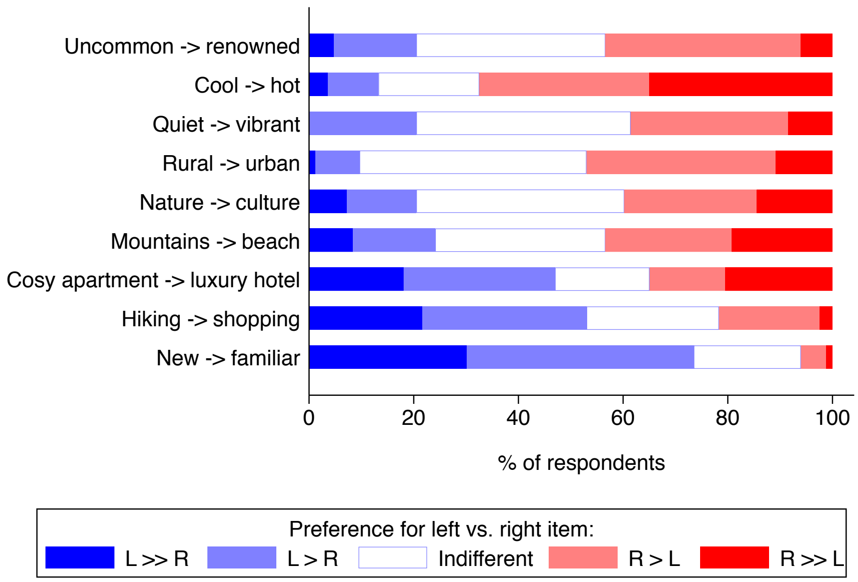
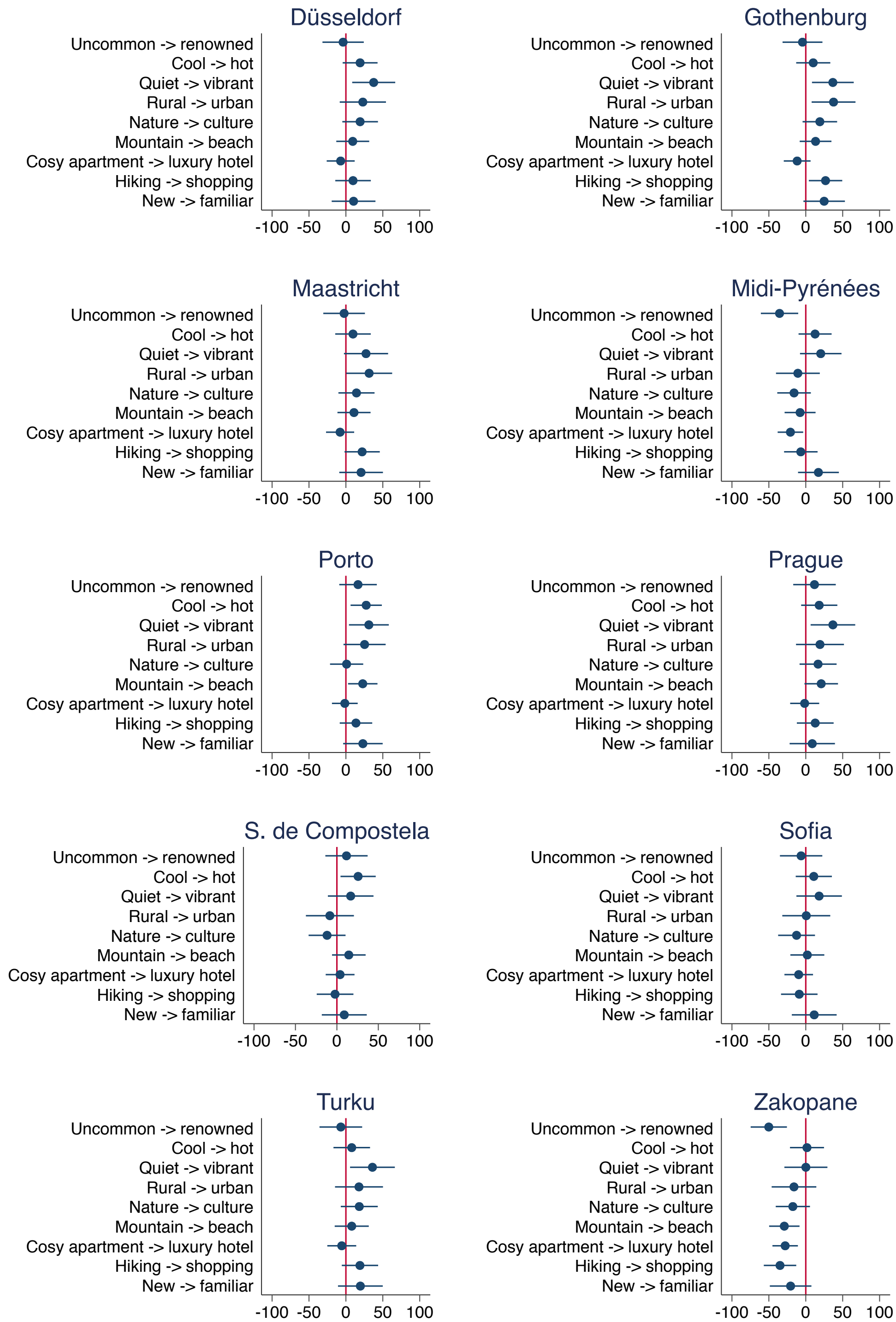


Figure E3: Relationship between valuations and trip preferences



Notes: Coefficients are from simple two-sided Tobit regressions of a respondent's valuation for trip k, $v_k \in [0, 500]$, on their location $c_j \in \{1, 2, 3, 4, 5\}$ along holiday criterion j; error bars are 95% CIs. For instance, a one-point increase on the scale "quiet → vibrant" is associated with a £37.67 increase (95% CI: [8.51, 66.82]) on average of a respondent's valuation for the trip to Düsseldorf.

F Alternative mechanisms

F.1 Ambiguity aversion and hedging

In this section, I provide information on the extent to which a hedging motive due to ambiguity aversion and/or a lack of trust in the experiment could explain the randomization decisions observed. First, I examine whether ambiguity attitudes measured in the domain of money are correlated with decisions to randomize over trips. Second, I report statistics on the level of trust in the experiment and its link to randomization decisions.

F.1.1 Ambiguity aversion in the domain of money

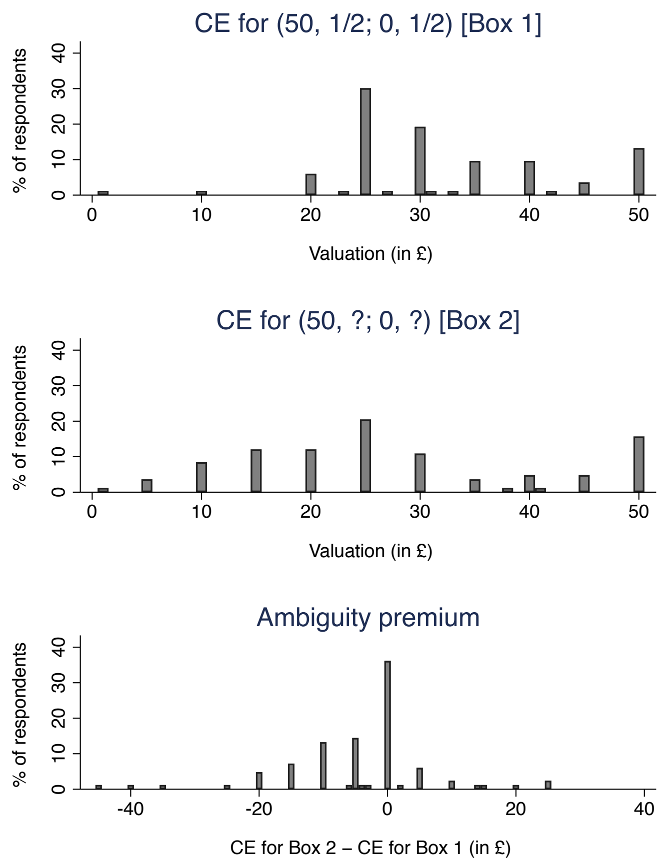
In Part 5 of the survey, respondents completed a standard valuation exercise to measure their preferences for one risky and one ambiguous monetary bet (Halevy, 2007; Chapman, Dean, Ortoleva, Snowberg, and Camerer, 2023). They were presented with 2 boxes, each containing 10 lottery tickets that could be either yellow or green:

- Box 1: contains 5 tickets of each color
- Box 2: contains an unknown number of yellow and green tickets

For each box, respondents were asked for their valuation of a bet that pays £50 if they correctly guessed the color of the ticket drawn drawn from the box. Valuations were elicited using the BDM mechanism by asking respondents for the minimum amount of money they would accept instead of taking the bet. Let P_j be the compensation chosen for Box $j \in \{1, 2\}$. Ambiguity aversion (love) is revealed from $P_1 > P_2$ ($P_1 < P_2$), with the ambiguity premium given by the difference $P_2 - P_1$. Before entering their valuation, respondents were asked to choose the ticket color to bet on; this was done in order to maximize trust. The composition of Box 2 was revealed right after subjects submitted their valuations, while the ticket color was only revealed at the very end of the survey, thus introducing meaningful temporal distance between the two lottery stages. Figure F1 shows the distribution of answers.

Unfortunately, a small fraction of respondents reported being confused about the term "£50 bet", as the instructions did not explicitly state that they would earn £0 in case they did not get the prize. This could explain the 13 (16) % of respondents who entered £50 as their valuation for Box 1 (2). However, to the extent that valuations for

Figure F1: Attitudes towards risk and ambiguity for money

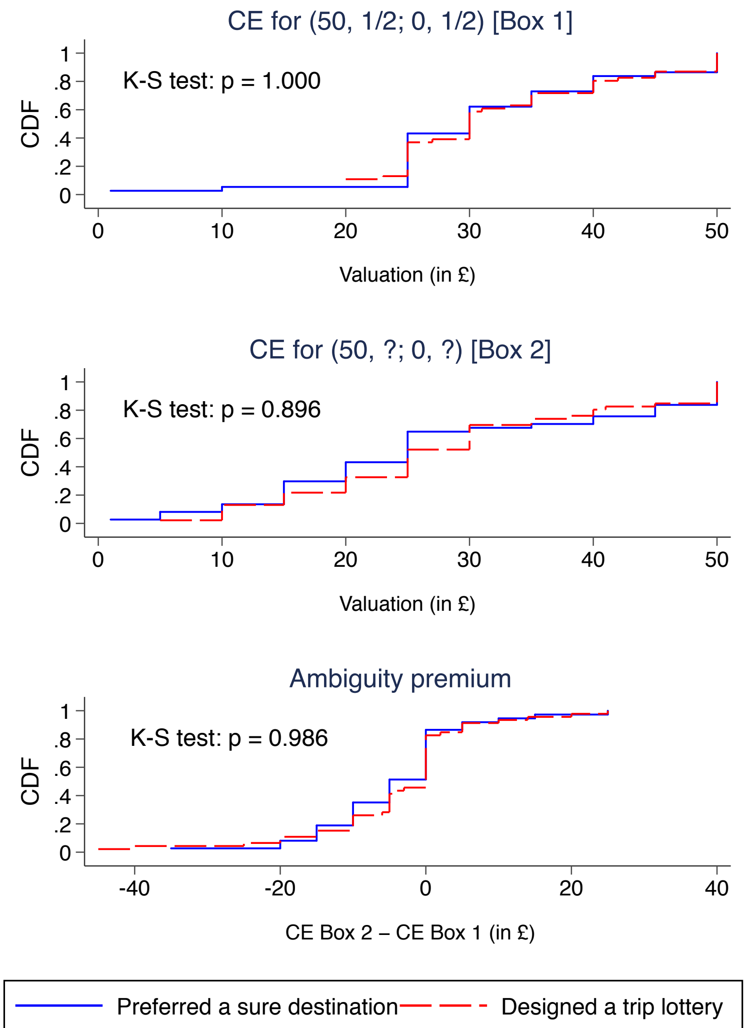


each box are affected by a similar amount of measurement error, taking the difference between the two should give an informative measure of the ambiguity premium.[17] Below I therefore focus my discussion on this measure. As the bottom panel of Figure F1 shows, there is overall ambiguity aversion. While the median respondent is ambiguity neutral (36% of respondents are), the average ambiguity premium is -£4.24 (95% CI: [-6.75, -1.74], p = 0.001), with 48% being classified as ambiguity averse and only 15% as ambiguity seeking. Figures F2 and F3 show how the ambiguity premium relates to preferences for surprise holiday trips.

While there is no relationship between ambiguity attitudes and propensity to build a trip lottery or SD violations in the binary decision problems, there is some suggestive evidence that respondents who built lotteries with more uncertainty or downside risk in them, who had a stronger preference for delay, or a higher valuation of the wildcard

[17] About 8% (7/83) of respondents entered £50 for each box.

Figure F2: Risk and ambiguity attitudes by favorite option

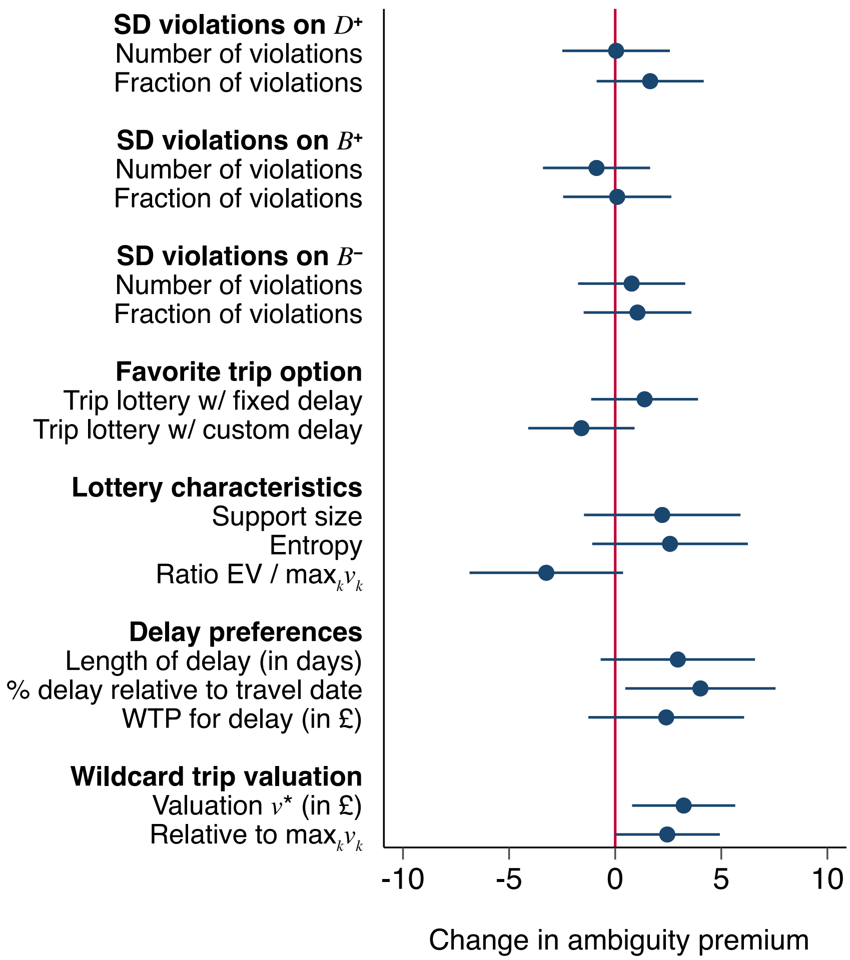


Notes: p-values from Kolmogorov tests of equality of distributions. "CE Box j" refers to the certainty equivalent or valuation $P_j \in [0, 50]$ entered for Box j. The ambiguity premium is measured as $P_2 - P_1$.

trip showed less aversion towards ambiguity. On the latter, the Pearson correlation coefficient between a respondent's wildcard trip valuation v^* and ambiguity premium is 0.28 ($p < 0.01$).[18] Given that the wildcard trip is ambiguous, this finding suggests some positive correlation in ambiguity attitudes across domains (trips and money).

[18] The correlation is 0.21 ($p = 0.05$) when correlating the ambiguity premium with the ratio $v^*/\max_k v_k$.

Figure F3: Correlates of the ambiguity premium



Notes: Coefficients are from simple linear regressions of `ambiguity premium (in £)` on the z-scores of the listed variables. Error bars are 95% confidence intervals. For instance, a one standard-deviation increase (relative to the sample mean) in the wildcard valuation v^* entered by a respondent is associated with a £3.2 increase (95% CI: [0.8, 5.7]) in the ambiguity premium. N = 83 (79) for the number (fraction) of SD violations on $\mathcal{D}^+$, $\mathcal{B}^+$ and $\mathcal{B}^-$; N = 83 for the variables under "Favorite trip option" and "Wildcard trip valuation"; N = 46 for the variables pertaining to "Lottery characteristics" and "Delay preferences".

F.1.2 Hedging due to mistrust

Figure F4: Trust in the experiment

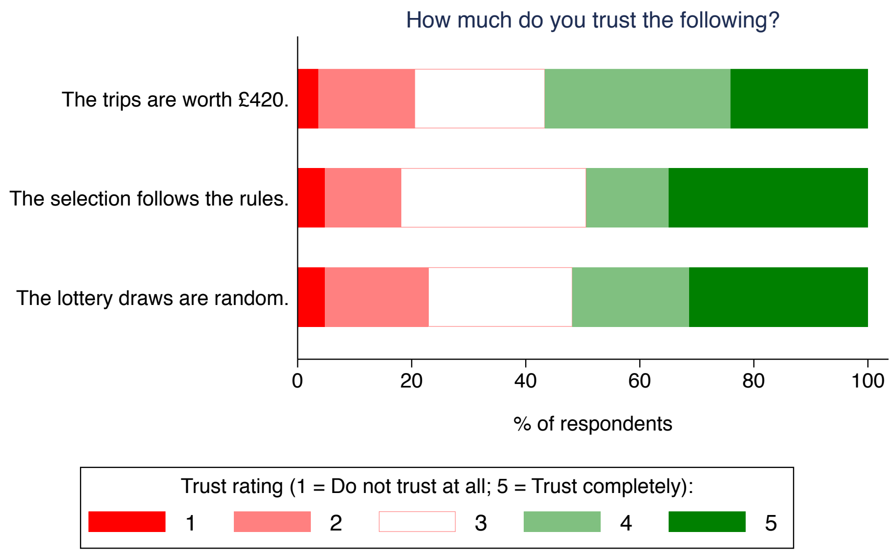
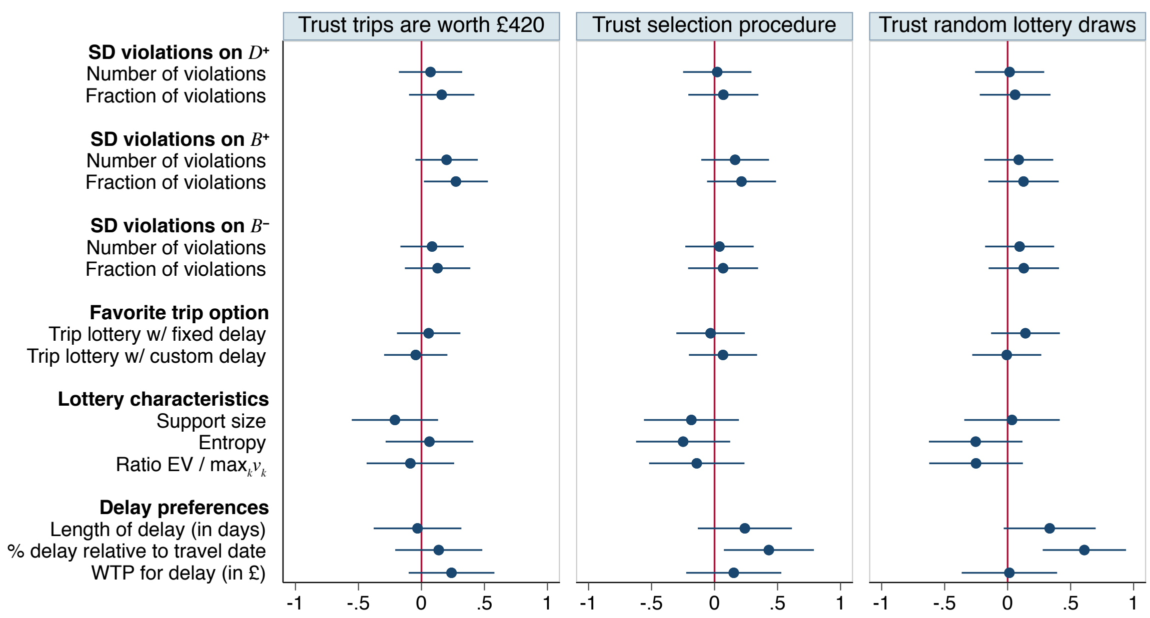


Figure F5: Relationship between trust and preference for surprise



Notes: Coefficients from univariate linear regressions of each trust rating (1-5) on the z-scores of the listed variables. Error bars are 95% CIs. N = 79 for the fraction of violations, N = 46 for variables under "Lottery characteristics" and "Delay preferences", and N = 83 for all the other variables.

F.2 Non-linear probability weighting

One possibility is that the DM has a distorted perception of the probability of each prize. These distortions could be purely cognitive and involuntary, or they could be motivated by a desire to think positively about the outcome i.e., if the DM derives utility from their beliefs. Below I examine what conditions on the probability-weighting function would be needed to generate the aggregate patterns observed in the experiment and discuss the realism of these conditions in light of the available evidence. Throughout, I maintain the following (minimal) assumptions on the probability-weighting function:

$$\pi'(p) > 0, \ \pi(0) = 0, \ \pi(1) = 1$$

To generate stochastic dominance violations, a necessary condition is that

$$\sum_{\{k \,:\, x_k \,\in\, \text{supp}(\mathbf{p})\}} \pi(p_k) \;\neq\; 1$$

Indeed, if the weights sum to 1, then $u(v_n) \leq \boldsymbol{\pi} \cdot \mathbf{u} \leq u(v_1)$ for any utility function u that is increasing in v and any set of weights $\boldsymbol{\pi} \in \Delta^n(X)$. Besides expected utility, this observation rules out rank-dependent probability weighting à la Quiggin (1982) (usually used in combination with expected utility i.e., the RDEU model) and the cumulative prospect theory (CPT) model of Tversky and Kahneman (1992), both of which assume that the weight $\bar{\pi}_k$ assigned to outcome x_k depends on the cumulative probability of obtaining at most x_k (not just on p_k) and satisfy by construction $\sum_{\{k \,:\, x_k \,\in\, \text{supp}(\mathbf{p})\}} \bar{\pi}_k \;=\; 1$.[19] Thus, those models cannot rationalize the data.

Next, consider the restrictions on $\pi(.)$ that are necessary to obtain the violations of P-MON observed in the experiment. In particular, recall that respondents were more likely to choose a simple lottery $(x_j, p; x_k, 1-p)$ over a sure option x_j such that $x_j \succ x_k$ (and, thus, $v_j > v_k$) when $p = 0.5$ compared to $p = 0.9$. Assuming non-linear

[19]For RDEU, $\bar{\pi}_k = \pi(\sum_{l \geq k}^{n} p_l) - \pi(\sum_{l > k}^{n} p_l)$, meaning that for a lottery $(x_j, p; x_k, 1-p)$ (with $x_j \succ x_k$), $\bar{\pi}_j = 1 - \pi(1-p)$ and $\bar{\pi}_k = \pi(1-p)$. A variant is used for CPT, with a different probability-weighting function for gains vs. losses.

probability weighting, the DM will choose $(x_j, p; x_k, 1-p)$ over x_j provided that

$$\pi(p)v_j + \pi(1-p)v_k > v_j \iff \frac{\pi(1-p)}{1-\pi(p)} > \frac{v_j}{v_k}$$

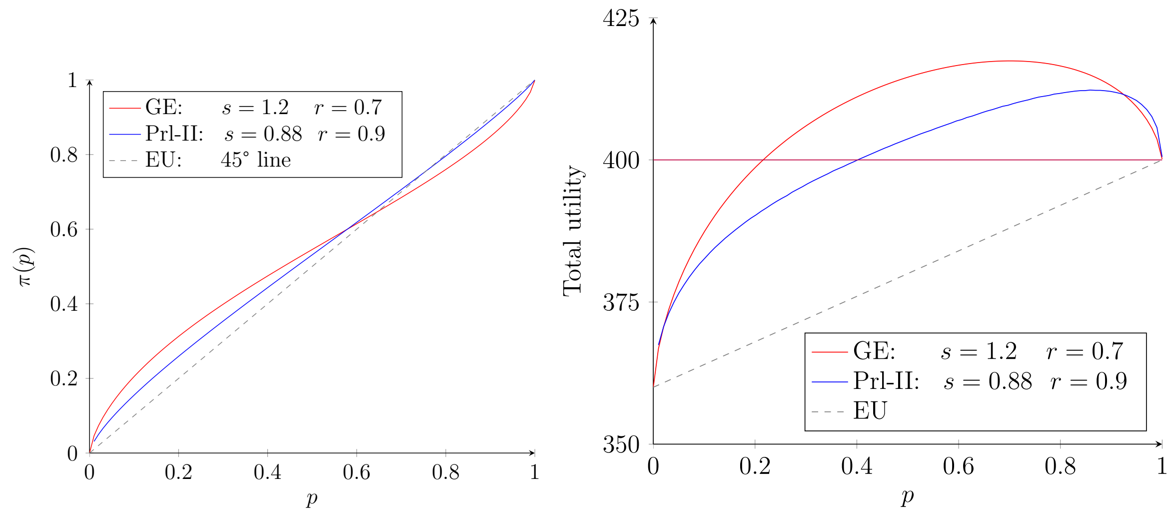
A necessary condition for P-MON to be violated when switching from $p = 0.9$ to $p = 0.5$ is $\pi(0.5) + \pi(0.5) > \max\{\pi(0.9) + \pi(0.1), 1\}$. The fact that subcertainty $\pi(p) + \pi(1-p) \leq 1$ must be violated at $p = 0.5$ contradicts the assumption of Kahneman and Tversky (1979), which has received empirical support for lotteries with monetary prizes.

Table F1 presents a list of the functional forms that have been used in the literature to estimate the degree of probability weighting of individual subjects (see Stott (2006)). Functions with a single parameter have insufficient flexibility to (realistically) accommodate the observed violations. In this category: (i) the functional form of Tversky-Kahneman (TK) was precisely chosen to generate subcertainty at all $p \in (0, 1)$; (ii) the one-parameter Prelec function (Prl-I) can avoid subcertainty at $p = 0.5$ for values of $r > 1$, but the probability-weighting function is then S-shaped in this parameter range (instead of inverse S-shaped), generating underweighting of small probabilities instead of overweighting (as is traditionally found). On the other hand, two-parameter functions (Wu-Gonzalez, Goldstein-Einhorn and Prelec II) can accommodate $\pi(0.5) + \pi(0.5) > 1$ while keeping $\pi(p)$ inverse S-shaped. The left panel of Figure F6 presents plots of the GE and Prl-II probability weighting functions with parameter values that generate $\pi(0.5) > 0.5$ and thus imply violations of stochas-

Table F1: Functional forms

Name	Abbreviation	Equation
Goldstein-Einhorn	GE	$\pi(p) = \frac{sp^r}{sp^r + (1-p)^r}$
Tversky-Kahneman	TK	$\pi(p) = \frac{p^r}{(p^r + (1-p)^r)^{\frac{1}{r}}}$
Wu-Gonzalez	WU	$\pi(p) = \frac{p^r}{(p^r + (1-p)^r)^s}$
Prelec I	Prl-I	$\pi(p) = e^{-(-\ln(p))^r}$
Prelec II	Prl-II	$\pi(p) = e^{-s(-\ln(p))^r}$

Figure F6: Parameter values that generate SD violations at $p = 0.5$



Notes: The left panel presents plots of the probability weighting function for two functional forms and parameter values that can lead to violations of stochastic dominance on $\mathcal{B}^+$ when $p = 0.5$. GE corresponds to $\pi(p) = \frac{sp^r}{sp^r+(1-p)^r}$ and Prl-II to $\pi(p) = e^{-s(-\ln(p))^r}$. The right panel plots the utility $\pi(p)v_j + \pi(1-p)v_k$ of a lottery $(x_j, p; x_k, 1-p)$ assuming $v_j = 400$ and $v_k = 360$ relative to the sure option $(v_j, 1)$ (pink horizontal line). The EU case corresponds to $\pi(p) = p$ (i.e., $s = 1$ and $r = 1$).

tic dominance on $\mathcal{B}^+$ (preference for randomization). For these probability-weighting functions, the right panel of Figure F6 plots the utility $\pi(p)v_j + \pi(1-p)v_k$ of a lottery $(x_j, p; x_k, 1-p)$ with $v_j = 400$ and $v_k = 360$ compared to the sure option $(x_j, 1)$. As the figure shows, although these parametric forms can accommodate SD violations at $p = 0.5$, they do not generate violations of P-MON when moving from $p = 0.5$ to $p = 0.9$: while the DM's utility from the lottery starts to decline at higher values of p, it remains higher than the utility of the sure option. In other words, if the DM violates stochastic dominance at some p, then they will continue to do so for any $p' \in (p, 1)$. To generate the observed violations of P-MON, one would therefore need to go beyond the parametric forms traditionally assumed in the literature.

Finally, it is worth considering what the implications of non-linear probability weighting might be for the correlation between SD violations on $\mathcal{B}^+$ vs. $\mathcal{B}^-$. Note that if subcertainty is preserved at $p = 0.1$ (i.e., $\pi(0.1) + \pi(0.9) < 1$), then the DM should also violate stochastic dominance on $\mathcal{B}^-$ (i.e., when comparing $(x_k, 1)$ to $(x_j, 0.1; x_k, 0.9)$ with $x_j \succ x_k$) if the difference in valuations is small enough. In fact, 31% of respondents preferred the sure option when $(j, k) \in \{(1, 2), (1, 3)\}$ and $p = 0.1$. However, I find a negative correlation between SD violations on $\mathcal{B}^+$ vs.

$\mathcal{B}^-$ in the aggregate, suggesting that these two kinds of SD violations correspond to different psychologies. In addition, if the probability weights are not rank-dependent, maintaining $\pi(0.1) + \pi(0.9) < 1$ implies $v_j > \pi(0.9)v_j + \pi(0.1)v_k$, meaning that respondents should not exhibit SD violations on $\mathcal{B}^+$ when $p = 0.9$. However, despite these violations being less common than for $p = 0.5$, they still occur about 30% of the time when $(j, k) \in \{(1, 2), (1, 3)\}$, which would require violating subcertainty.

In sum, while it is possible in principle to pick for each subject a probability-weighting function that will match the observed pattern of SD violations, the choice of function will have to substantially deviate from typical choices in the literature. Besides a consideration of the model's ability to fit the data, it is worth discussing its interpretation:

Cognitive noise: If probability distortions are purely cognitive i.e., the product of the DM's noisy perception of probabilities (e.g., due to a limited understanding of, or lack of familiarity with, the concept of probability), then those distortions should persist in any type of lottery that presents the same probability weights. In other words, according to this interpretation of non-linear probability weighting, the DM's distortion of a probability p should occur independently of the type of prize that this probability is attached to. In the experiment, the DM should therefore continue to violate stochastic dominance in decision problems where the trips are replaced by their valuations. Instead, violations almost completely disappear with money.

Wishful thinking: Another interpretation is that these probability distortions are the result of motivated beliefs about the outcomes that might realize e.g., to preserve optimism. If that is the case, one would expect the probability distortions to be rank-dependent, with the DM assigning a higher weight to outcomes ranked higher in their preference ordering (i.e., outcomes they enjoy more thinking about).[20] However, recent evidence casts doubt on rank dependence as a defining feature of probability weighting (Bernheim and Sprenger, 2020). In addition, one would need to come up with a model that would explain why the DM would prefer a lottery to the best outcome contained in it. Finally, the analysis of preferences for the wildcard trip suggests that optimism about the chances of getting a better outcome was not the main explanatory factor (see Section D.1).

[20] For a framework with this feature, see the model of Caplin and Leahy (2019) on wishful thinking. Note however that distorted probabilities sum to 1 in this model.

References

Bernheim, B. Douglas, and Charles Sprenger. 2020. "On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting". *Econometrica* 88.4, 1363–409. DOI: https://doi.org/10.3982/ECTA16646. [52]

Caplin, Andrew, and John V Leahy. 2019. "Wishful Thinking". Working Paper 25707. National Bureau of Economic Research. DOI: https://doi.org/10.3386/w25707. [52]

Chapman, Jonathan, Mark Dean, Pietro Ortoleva, Erik Snowberg, and Colin Camerer. 2023. "Econographics". *Journal of Political Economy Microeconomics* 1.1, 115–61. DOI: https://doi.org/10.1086/723044. [44]

Halevy, Yoram. 2007. "Ellsberg Revisited: An Experimental Study". *Econometrica* 75.2, 503–36. DOI: https://doi.org/10.1111/j.1468-0262.2006.00755.x. [44]

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47.2, 263–91. DOI: https://doi.org/10.2307/1914185. [50]

Quiggin, John. 1982. "A theory of anticipated utility". *Journal of Economic Behavior & Organization* 3.4, 323–43. DOI: https://doi.org/10.1016/0167-2681(82)90008-7. [49]

Stott, Henry P. 2006. "Cumulative prospect theory's functional menagerie". *Journal of Risk and Uncertainty* 32.2, 101–30. URL: https://doi.org/10.1007/s11166-006-8289-6. [50]

Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty". *Journal of Risk and Uncertainty* 5.4, 297–323. DOI: https://doi.org/10.1007/BF00122574. [49]