



Sleepiness, Choice Consistency, And Risk Preferences

By: Marco Castillo, **David L. Dickinson**, and Ragan Petrie

Abstract

We investigate the consistency and stability of individual risk preferences by manipulating cognitive resources. Participants are randomly assigned to an experiment session at a preferred time of day relative to their diurnal preference (circadian matched) or at a non-preferred time (circadian mismatched) and choose allocations between two risky assets [using the Choi et al. (Am Econ Rev 27(5):1921–1938, 2007), design]. We find that choices of circadian matched and mismatched subject are statistically similar in terms of satisfying basic requirements for preference consistency. However, mismatched subjects tend to choose riskier asset bundles.

Castillo, M., et al. (2017). "Sleepiness, choice consistency, and risk preferences." *Theory and Decision* 82(1): 41-73. <https://doi.org/10.1007/s11238-016-9559-7>. Publisher version of record available at: <https://link.springer.com/article/10.1007/s11238-016-9559-7>

Sleepiness, choice consistency, and risk preferences

Marco Castillo¹ · David L. Dickinson^{2,3,4} ·
Ragan Petrie¹

Abstract We investigate the consistency and stability of individual risk preferences by manipulating cognitive resources. Participants are randomly assigned to an experiment session at a preferred time of day relative to their diurnal preference (circadian matched) or at a non-preferred time (circadian mismatched) and choose allocations between two risky assets [using the Choi et al. (*Am Econ Rev* 27(5):1921–1938, 2007), design]. We find that choices of circadian matched and mismatched subject are statistically similar in terms of satisfying basic requirements for preference consistency. However, mismatched subjects tend to choose riskier asset bundles.

Keywords Risk aversion · Rationality · Cognitive resource depletion · Choice consistency · Sleepiness

B Ragan Petrie
rpetrie@tamu.edu

Marco Castillo
marco.castillo@tamu.edu

David L. Dickinson
dickinsondl@apstate.edu

¹ Department of Economics, Texas A & M University, 4228 TAMU, College Station, TX 77843, USA

² Center for Economic Research and Policy (CERPA), Department of Economics, Appalachian State University, Boone, NC, USA

³ IZA, Bonn, Germany

⁴ ESI, Chapman University, Orange, CA, USA

1 Introduction

There is growing evidence that individual risk attitudes, as measured by economic experiments, vary across people and circumstances. These include life-cycle changes, traumatic personal or family experiences (Voors et al. 2012; Malmendier and Nagel 2011; Callen et al. 2014), physical conditions (Garbarino et al. 2011; Wozniak et al. 2014), priming and framing (Benjamin et al. 2010), cognitive ability (Dohmen et al. 2010; Burks et al. 2009; Benjamin et al. 2013), the different way in which some may bracket choices (Read et al. 1999) and by ones' genetic makeup (Cesarini et al. 2009).

In this paper, we examine if temporary challenges to cognitive functioning yield choices that are consistent with rationality or whether differences in decisions are due to lapses in rational behavior broadly defined (e.g., choosing randomly). In particular, we investigate whether a sleepiness manipulation through circadian mismatch, which is shown to be associated with impairment of cognitive abilities (Bodenhausen 1990; Kruglanski and Pierro 2008; Dickinson and McElroy 2012), produces changes in preferences while maintaining consistency of behavior. A common assumption in standard economic models (e.g., Arrow-Debreu state-dependent preferences model, Mas-Collel et al. 1995), as well as behavioral models (Tversky and Kahneman 1992; Koszegi and Rabin 2007; Becker and Murphy 1988), is that changes in preferences can occur without the loss of rationality. Our experimental design provides empirical evidence to investigate this assumption.

Circadian timing of decisions is a natural environment in which to test the stability and consistency of preferences. First, sleepiness has been widely studied in the sciences, and its effects on performance in many domains are well documented and understood.¹ Second, it is a physical condition commonly experienced by most people at some, or many, period(s) of their lives. Because of this, circadian mismatch, compared to other ways to temporarily deplete cognitive resources, is a manipulation that is less likely to generate inconsistencies in behavior due to learning or adaptation to the circadian mismatch. This is important because such learning would confound an examination of preference consistency across states.

The optimal time-of-day for alertness varies across individuals because of differences in morning or evening diurnal preferences. While time-of-day preference may tend towards morning time as individuals age, it is considered largely independent of gender (see Paine et al. 2006). We, therefore, recruit only young men and women who are validated morning- and evening-type individuals.² Morning- versus evening-type preference individuals are known to have daily alertness/sleepiness cycles that peak at different times-of-day (see Smith et al. 2002), independent of the sleep inertia that

¹ Total sleep deprivation studies are a more common approach to studying how sleepiness affects performance and decisions. However, circadian mismatch is a milder, and arguably more externally valid, way to examine sleepiness of the sort commonly experienced by real-world decision makers. Having to be awake at a time of day that is not optimal is more common and may affect performance in a less extreme manner than it would following 24 or more hours of total sleep deprivation in a relatively foreign sleep laboratory environment.

² Our sample is 95.9 % comprising young adults between 18 and 25 years old (99 % between 18 and 36 years old).

builds from cumulative hours awake.³ Importantly, results from this type of environment should be relevant to policy. Understanding whether risky choice decisions while sleepy are rational or not and if preferences change could help inform the design of institutions and policies.

Our research protocol is designed to reduce the issue of selection, and allow interpretation of our results due to temporary cognitive resource depletion. Specifically, once we identify morning- and evening-type subjects they are randomly assigned to a session (i.e., morning or evening session) and not allowed to select their own session time. This minimizes the problem of subjects selecting into session by time of day preference.

To identify validated morning- and evening-type individuals, we invite by email all students at two large universities to complete an online survey (see Appendix 2A for full set of pre-screen survey questions). Within the survey, there is a validated instrument to measure diurnal preference (Adan and Almiral 1991). The data are then used to identify two classic diurnal preference groups: those who are naturally most alert in the morning and those who are naturally most alert in the evening. These morning types and evening types were then randomly assigned to one of two session times: early morning or late night. This produced two treatments, participants who were “matched” in terms of their circadian rhythm (e.g., morning type in a morning session and evening type in an evening session) and circadian “mismatched” (e.g., morning type in an evening session and evening type in a morning session). Participants assigned to one session time were not allowed to switch to the other. Compliance with session assignment was voluntary and, importantly, we find no significant differences in compliance across treatment conditions. Participants were allowed to take all the time they needed to make their decisions, and this was done to allow participants the opportunity to express their preferences unconstrained by time.

Our results show a significant treatment effect on risk decisions, and as mentioned above, this is not due to selection or compliance across treatments. Circadian mismatched participants have higher certainty equivalents for different risky asset bundles, indicating they are less risk-averse. Also, the variance in risky asset investment is larger for these subjects, showing a tolerance for more variability in payments.

While the manipulation clearly worked as designed and affected preferences, it did not alter the likelihood a subject behaved rationally. Adherence to the generalized axiom of revealed preference (GARP) is identical between mismatched and matched participants.⁴ As a result, the estimated behavioral differences result cannot be attributed to an increase in noisiness of the data following circadian mismatch because an increase in noise would manifest in increased violations of choice consistency of

³ It is also true that validated morning-type individuals are much more rare in young adult populations, such as our college student samples (Chelminski et al. 1997, 2000). Therefore, we recruit a higher proportion of the morning-types available in our student populations in order to achieve roughly equal numbers of morning-type and evening-type subjects in our final sample.

⁴ Similarly, deviations from expected utility theory (EU) and rank-dependent utility theory with a concave utility function are also statistically similar across groups. These results are presented in Appendix 2C. We also examine whether subjects violate payoff dominance in making choices and, while both groups display some violations, there is no difference in violations between groups. In our data, consistency with rationality is robust.

one sort or another. All of this suggests that preferences can be altered without altering adherence to rational behavior.

Our paper contributes to the literature by showing that a slight manipulation of physical conditions, which has been shown to produce a temporary challenge to cognition, produces changes in risk attitudes without producing a breakdown of rationality. The reader should note that our results, whereby we present a temporary cognitive resource challenge, are not meant to be directly compared to studies examining the impact of *permanent* cognitive abilities on risk taking (e.g., [Dohmen et al. 2010](#); [Burks et al. 2009](#); [Frederick 2005](#)).⁵ Our manipulation leads to higher certainty equivalents for mismatched subjects, indicating an increased preference for monetary risk when sleepy. Our experiment was not designed to identify the mechanism causing these effects. However, our results do show that the relationship between sleepiness and preferences is causal.

To our knowledge, this is the first paper to show that changes in preferences can occur without loss of rationality. The closest paper to ours is that of [Burghart et al. \(2013\)](#), who examine behavioral effects of alcohol intake but conclude that it does not impair rational decision-making. In our data, we do identify a change in risk preference in spite of no difference in rationality.⁶

In Sect. 2, we describe the experimental design and the cognitive resource manipulation. Section 3 reports our results, first by confirming that our manipulation worked, and then examining rationality and choice behavior in the risk task. Finally, Sect. 4 concludes.

2 Experimental design

2.1 The risky choice experiment environment

We follow the design of [Choi et al. \(2007\)](#) for the risky choice task, which generates a rich set of individual-level data.⁷ In each decision round, subjects are asked to allocate tokens between two different accounts: X and Y. Tokens in account X only generate a payoff for the subject if account X is randomly chosen by the computer at the end of that decision round. Similarly, tokens in account Y only pay if account Y is randomly selected. We implement the “symmetric” treatment design of [Choi et al. \(2007\)](#) with a common knowledge 50 % probability that either account X or Y will be chosen. Because accounts X and Y are equally likely to be chosen, the accounts are perfect

⁵ This other stream of research has shown that individuals with higher levels of permanent cognitive ability display an increased propensity to take monetary risk. While interesting, this is fundamentally distinct from our question of how temporary depletion of available cognitive resources will impact risky choice and/or rationality.

⁶ Our design includes 50 decisions per subject and provides more statistical power, which increases our confidence in the finding that rationality is not altered by our manipulation. Also, in our design, participants were randomly assigned to treatments (circadian mismatched or matched), and this reduces the possibility that participants self-select into the experiment based on certain characteristics (e.g., rationality).

⁷ We are grateful to Sachar Kariv for providing us with the code for the experiment task. Subject instructions for the experiment are included in Appendix 1.

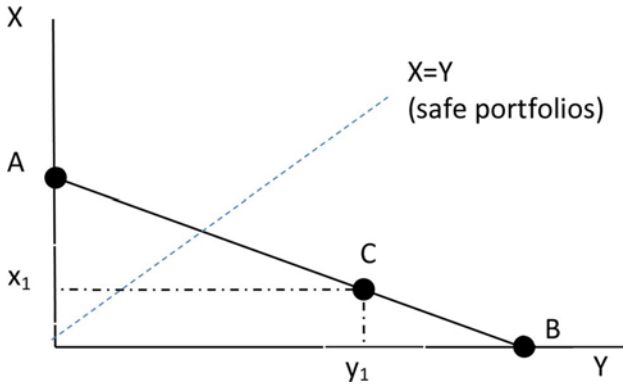


Fig. 1 Sample stimuli

substitutes for one another. Figure 1 shows a sample stimulus where the subject makes an allocation choice on a computer interface using a mouse-driven pointer to drag point C along the line AB to their desired choice location (including the endpoint locations, if desired). An allocation such as point A or point B is a risky choice with all tokens placed in one account. Thus, the subject would only receive a payoff if the computer randomly selects the account where all the tokens are allocated. An intermediate allocation of tokens, such as point C in Fig. 1, places some tokens in each account, which guarantees the subject a smaller, but sure, payoff in both states of the world. A choice along the $X = Y$ line in Fig. 1 is a perfectly safe portfolio that guarantees the same payoff no matter which state of the world applies.

The experiment consists of 50 decision rounds (i.e., 50 different stimuli) where the slope and intercept of the AB line are randomly determined for each stimulus.⁸ After all 50 rounds are complete, one round is randomly selected for payment, and each round has an equal probability of being chosen. The randomly selected payoff-round, the computer's random selection of account X or Y, and the subject's allocation decision for that round determine the subject's payoffs.

Subjects complete a questionnaire to assess their current level of sleepiness before starting the experiment and also upon completion of the decision task. They also respond to several self-report sleep questions and complete a cognitive reflection task (CRT) after the decision task (see Appendix 2B).

2.2 The cognitive resource manipulation

We use a circadian match/mismatch protocol to represent a temporary challenge to cognitive resources. While there may be other ways to temporarily deplete resources,

⁸ For the first eight sessions, the X and Y intercepts were constrained to lie in the [50, 100] interval (see Fig. 1). For the latter eight sessions, in order to generate more extreme relative prices, the budgets were chosen among the set of lines that intersect both axes below 100 and intersect at least one axis at or above the 50-token level. The initial starting point for the mouse-pointer along each budget line was also randomly determined. See Choi et al. (2007) for full details.

our method has broad applicability to circumstances encountered in daily life, has been previously used and validated in the literature, and is relatively easy to administer.

Previous research has shown that single-vehicle accidents increase at times of the day where the typical circadian rhythm dictates sleepiness due to natural release of melatonin (Coren 1996). In controlled experimental settings, researchers have found that sleep-deprived individuals take on more risk than well-rested individuals when choosing between risky lotteries (McKenna et al. 2007). A more recent study shows that increased risk taking when totally sleep deprived may be specific to male subjects (Ferrara et al. 2015). In a different risky choice task, Venkatraman et al. (2007) found neural effects in sleep-deprived subjects even in the absence of behavioral effects. Though sleep loss and circadian timing may both contribute to depleted cognitive resources with symptomatic sleepiness, a 24-h total sleep deprivation protocol likely depletes cognitive resources to a greater extent than what individuals commonly experience on a daily basis and may, therefore, not be applicable to a large segment of the population. Decision making at sub-optimal times of the day is more externally valid, and hence motivates our choice of the circadian mismatch protocol.

Explicit circadian mismatch protocols like we propose have been used in behavioral research to some extent, but this area is relatively unexplored. Bodenhause (1990) showed that individuals are more likely to use stereotypes in making judgments when at circadian mismatched times, and Kruglanski and Pierro (2008) reported an increased use of the psychological transference effect among subjects tested at circadian mismatched times. Dickinson and McElroy (2010) and Dickinson and McElroy (2012) used two distinct protocols to manipulate the circadian timing of decision in guessing games, and find that choices made at circadian mismatched times generally produce outcomes farther from the predicted Nash equilibrium. Though limited, the extant literature on circadian mismatch effects is consistent with the hypothesis that circadian mismatch alters decision making in a way consistent with cognitive resource depletion.⁹

To implement the circadian mismatch protocol for our current study, we first administer a large-scale online survey at two large universities. The objective of the survey is to generate a database of individuals for whom we have a validated measure of their diurnal preference, which is assessed in the survey using the short form of the morningness–eveningness questionnaire, henceforth rMEQ (Adan and Almiral 1991). The rMEQ classifies individuals on a scale of 4–25, with morning-types having rMEQ score from 18 to 25 and evening-types having rMEQ score from 4 to 11. While this diurnal preference measure is based on self-reports of the subjects, it has been validated against physiological data on oral temperatures (see Horne and Östberg 1976) and is a standard tool in circadian research.

From our database, we recruit morning-types and evening-types, who we had randomly assigned, *ex ante*, to participate in either a morning (7:30 a.m.) or an evening (10:00 p.m.) experiment session. This resulted in 57 % of our sample being circadian

⁹ To be more specific, depletion of cognitive resources would disproportionately affect executive function. The behavioral effects reported in the aforementioned studies either involve increased reliance on heuristics (i.e., stereotyping and transference effect) or decreased ability to engage in strategic reasoning (i.e., the guessing game), both of which are consistent with reduced engagement of deliberate thought regions of the brain that rely on fully intact cognition.

Table 1 Sample size per design cell

	Morning session	Evening session
Morning-type	32 + 22 = 54	29 + 23 = 52
Evening-type	24 + 11 = 35	33 + 28 = 61
	Total sample size = 202	
	Matched obs = 115, Mismatched obs = 87	

(Site 1 + Site 2 = total sample size), Mismatched (depleted cognitive resource) cells located in off-diagonal

matched for the risky choice experiment.¹⁰ If a subject could not participate in the randomly assigned time-slot (morning or evening), the subject was *not* allowed the option of the alternative time-slot—an alternative time-slot was not even mentioned in recruitment. This aspect of our design eliminates selection into treatments and allows for a more causal interpretation of cognitive resource depletion on outcomes. Importantly, we find no evidence of selection in show-up rates across the matched and mismatched subjects. The proportion of subjects who actually showed up for the session they signed up for is not significantly different across our matched and mismatched subjects (the *p* value of a Chi-square test of difference in distribution is 0.516).¹¹

In total, 202 subjects (102 men, 100 women) participated in this study. Table 1 shows the distribution of our sample across experiment locations for each design cell.¹² The experiment sessions lasted just over an hour, and included the risky choice task administration as well as a few short survey instruments to elicit self-reported measures of recent sleep habits (shown in Appendix 2B). Average subject payoffs were \$22.56 (s.d. \$9.61), which includes a \$5 show up fee.

3 Results

3.1 Summary statistics and manipulation check

Table 2 shows relevant summary statistics of our sample. Pre-experiment survey data refer to responses from the online sleep survey administered as a way of building our database of morning-type and evening-type subjects. Pre-experiment survey responses would have been given several days to several weeks before the decisions experiments.

¹⁰ Due to the rarity of true morning-type subjects—less than 10 % in young adult populations are morning-types (see Chelminski et al. 2000)—we extend our rMEQ cutoff to include rMEQ scores of 15–17. To compensate, we only recruit the more extreme (and still abundant) evening-type subjects with rMEQ scores from 4 to 10. In this way, our sample is still drawn from the tails of the rMEQ distribution and eliminates the same amount of support from the non-tail portion of the rMEQ distribution compared to if we had used the traditional morning-type cutoff (rMEQ = 18) but included non-extreme evening types (rMEQ = 11) in our sample.

¹¹ One-hundred and thirty-seven matched subjects signed up for experimental sessions, and 115 showed up. One-hundred and ten mismatched subjects signed up for experimental sessions, and 87 showed up.

¹² Men and women were equally distributed across the matched and mismatched groups (matched: 62 men, 53 women; Mismatched: 40 men, 47 women) and there is no significant difference in the distributions (Fisher's exact test, *p* value = 0.320)

Table 2 Descriptive statistics of sample

	Circadian matched (<i>n</i> = 115)	Circadian mismatched (<i>n</i> = 87)	<i>p</i> value for diff in means test (<i>t</i> test)
Pre-experiment survey			
Avg hours of sleep	6.8	6.8	0.9195
Hours slept last night	6.8	6.8	0.8048
Optimal hours of sleep	8.0	7.9	0.8357
Epworth trait-level sleepiness score (0–24, higher numbers indicate chronic fatigue)	7.6	7.4	0.5855
Post-experiment survey			
Avg hours of sleep over last 7 days	7.1	6.7	0.0382
Hours slept last night	6.7	6.2	0.0515
Optimal hours of sleep	8.0	8.0	0.8150
Epworth trait-level sleepiness score (0–24, higher numbers indicate chronic fatigue)	8.0	8.4	0.3600
Cognitive Reflection Task (CRT) score (0–3)	1.56	1.68	0.4677
Manipulation checks			
Avg state-level sleepiness rating during experiment (Karolinska scale) (1=extremely alert, 9=extremely sleepy)	4.5	5.7	0.0000
Avg decision response time in seconds (std dev of mean in parentheses)	9.4 (4.07)	10.2 (5.00)	0.2157
Distribution of response times			Kolmogorov– Smirnov test for difference in distributions <i>p</i> value = 0.002

Table 2 includes summary statistics for the same questions asked in the pre-experiment survey and after subjects had completed the risky choice task but before payments were revealed (section labeled “post-experiment”). These include self-reports of the subjects’ average nightly sleep over the 7-days prior to the experiment, average sleep the night prior to the experiment, one’s subjective optimal hours of nightly sleep required for peak performance, and the Epworth Sleepiness scale used commonly in sleep research (a measure of trait-level sleepiness, or chronic fatigue). As can be seen in the top panel, there are no significant differences in any of these descriptive measures between the circadian matched and mismatched subjects in the pre-experiment

survey.

In the post-experiment survey data in Table 2, we find that circadian mismatched subjects report being significantly more sleep restricted than circadian matched sub-

jects. We attribute this to factors particular to the morning sessions, when sleep level difference likely results from circadian matched subjects following a more typical schedule the evening prior to the experiment compared to mismatched (evening-type) subjects. Specifically, evening-type subjects mismatched to the morning session more likely restricted their sleep to wake early for the experiment. Truncation of the previous night's sleep due to our experiment schedule is not an issue with scheduling of circadian matched subjects (e.g., a morning-type subject in the morning session more likely had a typical level of sleep before the experiment). This same issue is not present in our evening sessions, whether both morning-types and evening-types could have slept; however, they pleased the night before. This explanation is consistent with our data that show the lowest average nightly sleep levels among evening-type subjects in the morning sessions ($n = 36$ subjects: mean nightly sleep = 5.26 ± 1.28 h). Finally, we note that our mismatched subjects reporting significantly more sleep restriction only serve to amplify the adverse cognitive resource state we intended to manipulate. It also highlights the complexity of trying to separately manipulate circadian effects from sleep level effects in an experimental design of this nature.

Because our objective is to introduce a randomized assignment of cognitive resource availability, we present evidence that our circadian manipulation was successful. The manipulation checks in Table 2 reveal that circadian mismatched subjects, who presumably have depleted cognitive resources, report significantly higher state-level sleepiness. There is no significant difference in the score on the CRT task across the two groups. Matched and mismatched subjects do not differ in average length of time to make decisions. However, the median response time and distribution of time to make decisions are different.¹³ We also confirm that, on average, circadian matched and mismatched subjects faced similar menus—average intercepts of randomly generated budget lines in Fig. 1 do not significantly differ in mean or variance across.¹⁴

In sum, we validate that our sleepiness manipulation worked and the decision environment for our matched and mismatched subjects was the same.

¹³ There is no a priori reason why sleepy subjects would be expected to make faster or slower decisions. While some dual systems frameworks identify longer response times with more deliberate thinking, in the case of sleepiness, longer response times might also reflect difficulty focusing on the decisions. A recent theoretical framework of response times (Achtziger and Alós-Ferrer 2014) allows an initial stage of decision making where individuals first decide whether to use automatic or deliberative thinking. As such, overall response times are not suitable for identifying whether individuals are using more automatic or deliberative thought processes in our design, which was never intended to produce such informative response time data. Mismatched subjects are slower in responding when relative prices for the two assets are close to one. This could reflect that sleepiness makes it more difficult to choose between similar options or could reflect some kind of indifference that is exacerbated for sleepy subjects (Krajbich et al. 2015, document slower response times for decisions over options for which a subject may be indifferent). The median response time of matched subjects is 7 s and the median response time of mismatched subjects is 8 s.

¹⁴ A Kolmogorov–Smirnov test for difference in the distribution of the maximum intercept (max amount a person can put in one asset) that matched and mismatched subjects saw shows no significant difference. The test was run separately for the first 8 sessions and the last 8 sessions because these differed in the budget generation process (see footnote 8). The p value for the first 8 sessions is 0.674 and for the last 8 sessions is 0.868. The same distribution test was also run for the slope (relative prices) for the two sets of sessions, and there is no significant difference (p value = 0.77 and 0.218).

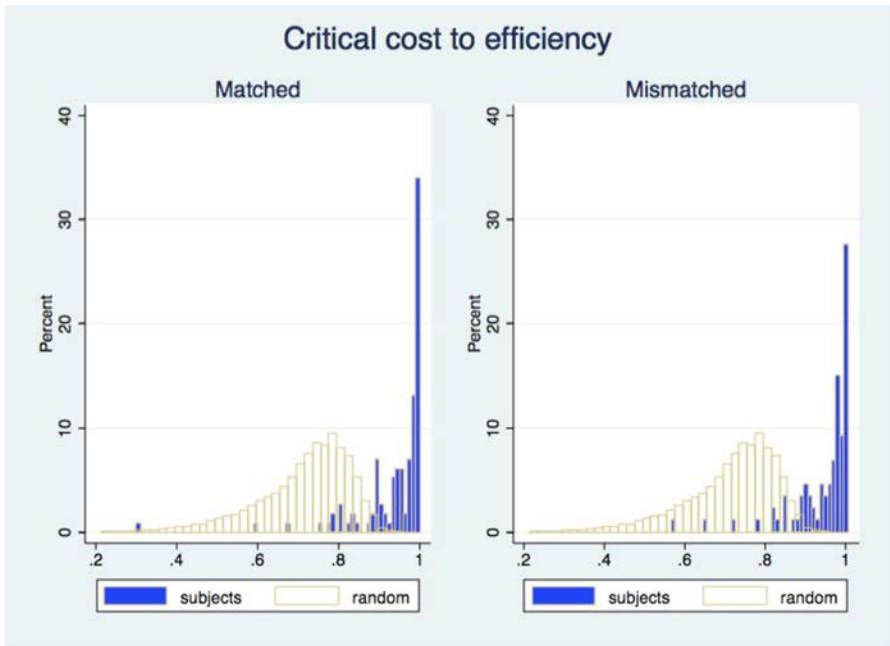


Fig. 2 Distribution of critical cost to efficiency Index (CCEI) for matched and mismatched subjects (*blue bars*) and for random choices (*yellow bars*)

3.2 Consistency of behavior

We look first at rationality and then choices in the risk task. We test for consistency of choices with rationality for matched and mismatched subjects using several measures of rational behavior. Specifically, we test if subjects satisfy the Generalized Axiom of Revealed Preferences (GARP: [Afriat 1972](#); [Varian 1983](#)) and payoff dominance.¹⁵ This examination of choice rationality provides a test of whether there is an increase in the noise of the choice data when sleepy. Noisier choice data would result in increased violations of GARP and payoff dominance. As we will see below, this is not the case.

Figure 2 presents the distribution of the Critical Cost to Efficiency Index, CCEI ([Afriat 1972](#)) for matched and mismatched subjects. This index measures how much budget constraints would need to be adjusted to eliminate all violations of GARP. CCEI values equal to one indicate “rational” behavior in the sense that no adjustment needs to be made to a budget constraint in order to eliminate a violation of GARP. As a point of reference as to how the CCEI measures rational choices, the figure also shows the distribution of CCEI for random choices. This distribution is shown in light bars, and the distribution for subject choices is shown in blue bars. The two

¹⁵ Additional tests of rationality (expected utility and rank-dependent utility with a concave utility function) produce the same results. More information and technical details on the tests we performed are given in Appendix 2C. The main conclusion is that we find no significance difference between matched and mismatched subjects using a variety of definitions of rationality or adherence to common utility preference frameworks.

distributions are significantly different (Chi-square test of difference in distribution p value=0.000), which shows that subject choices in the experiment violate GARP less often than would be expected from random choices.¹⁶

Looking at the CCEI's of subjects, Fig. 2 shows that 13 % of matched and mismatched subjects satisfied GARP without having to modify any budget (CCEI = 1). An additional 21 % of the matched subjects and 22 % of the mismatched subjects require a small change in the budgets to satisfy GARP (CCEI above 0.999). All told, 78 % of the matched subjects and 83 % of the mismatch subjects have indices strictly above 0.9. Indeed, the distribution of the CCEI is not significantly different between these two groups (rank sum test of difference in distributions p value = 0.7805). That is, slight cognitive resource depletion does not cause an increase in the distant to rationality, as measured by the CCEI.

Adherence to payoff dominance makes clear predictions in the context of this choice task. Namely, all choices on the short side (e.g., above the $X = Y$ line in Fig. 1) of the budget line would violate payoff dominance in the sense that one is strictly worse off by choosing such bundles compared to bundles on the long side of the budget line. This is true due to the fact that in our design asset X and asset Y are perfect substitutes, and so a subject choosing any asset bundle that is not on the $X = Y$ (safe bundle) line should always choose the bundle with the higher expected payoff. Similar reasoning explains why bundle A should never be chosen over bundle B in Fig. 1. Both bundles involve a 50 % chance of receiving a zero payoff, but the possible 50 % payout is higher with bundle B. Our data set provides ample observations to examine violations of payoff dominance in our two experimental groups.

Table 3 presents the contingency table of violations of payoff dominance for the steep versus flat budget constraints (the rows), and it does so for the case of the entire data set as well as subsets of the data for which relative prices are quite close (the columns). The importance of investigating relatively close price ratios is because violations of dominance are less severe in those cases, at least in terms of the magnitude of the expected payoff loss. Observing significant differences across matched and mismatched participants for close price ratios would be a strong test for noisy decision making. Fisher's exact tests are performed for each column represented in Table 3.

Two things are clear from Table 3. First, there are a significant number of violations of payoff dominance. It is violated roughly 1/3 of the time for the set of all budget constraints and, not surprisingly, dominance is violated more frequently as we constrain the data to the set of relative prices closer to one (i.e., the

and $\ln \frac{P_x}{P_y} < 0.10$ and $\ln \frac{P_x}{P_y} < 0.05$ subsamples). That is, as expected, violations increase as the

cost of a violation decreases. The second observation from Table 3 is that there is no

¹⁶ We only consider positive prices in all the analyses in the paper. Out of the 10,100 choices made in the experiment (50 choices per subject \times 202 subjects), 10,094 were with strictly positive prices as our failure to constrain the random price choice in the parameterization of the experiment inadvertently led to

6 instances of zero-priced assets.

¹⁷ We use the log because it gives the percentage change of relative prices [e.g., $\log(1.10) \approx 0.10$].

Table 3 Consistency versus Violations of Payoff Dominance: Matched vs. Mismatched subject Fisher tests of proportions for each category (number listed, proportion of sample in parenthesis)

Category	These columns examine subsets of budget lines reflecting smaller differences between the prices of assets X and Y (i.e., budget lines with slopes closer to -1)		
	All $\ln \frac{P_x}{P_Y}$ ($n = 10,100$)	$\left \ln \frac{P_x}{P_Y} \right < 0.10$ ($n = 2203$)	$\left \ln \frac{P_x}{P_Y} \right < 0.05$ ($n = 1156$)
Relatively cheap X Dominance consistent choice	$N_{\text{Matched}} = 2128$ (73.46 %)	$N_{\text{Matched}} = 371$ (57.61 %)	$N_{\text{Matched}} = 193$ (53.76 %)
	$N_{\text{Mismatched}} = 1647$ (74.42%)	$N_{\text{Mismatched}} = 295$ (59.84%)	$N_{\text{Mismatched}} = 143$ (55.21%)
Relatively cheap Y Dominance consistent choice	$N_{\text{Matched}} = 2208$ (77.39 %)	$N_{\text{Matched}} = 393$ (63.80 %)	$N_{\text{Matched}} = 204$ (63.75 %)
	$N_{\text{Mismatched}} = 1687$ (78.94 %)	$N_{\text{Mismatched}} = 311$ (69.11 %)	$N_{\text{Mismatched}} = 140$ (64.22 %)
Relatively cheap X Dominance violated choice	$N_{\text{Matched}} = 769$ (26.54 %)	$N_{\text{Matched}} = 273$ (42.39 %)	$N_{\text{Matched}} = 166$ (46.24 %)
	$N_{\text{Mismatched}} = 566$ (25.58 %)	$N_{\text{Mismatched}} = 198$ (40.16 %)	$N_{\text{Mismatched}} = 116$ (44.79 %)
Relatively cheap Y Dominance violated choice	$N_{\text{Matched}} = 645$ (22.61 %)	$N_{\text{Matched}} = 223$ (36.20 %)	$N_{\text{Matched}} = 116$ (36.25 %)
	$N_{\text{Mismatched}} = 450$ (21.06 %)	$N_{\text{Mismatched}} = 139$ (30.89 %)	$N_{\text{Mismatched}} = 78$ (35.78 %)
Fisher's exact test	0.465	0.247	0.948

significant difference in propensity to violate dominance between circadian matched and mismatched participants (see Fisher's test results in bottom row of Table 3). Violations of dominance are actually uniformly *lower* among the circadian mismatched group, but the Fisher's exact tests show no statistically significant differences. These results complement the rather extensive examination of consistency with other definitions of rationality (see Appendix 2C), but all tests in this section point to the same result.

Result 1 Cognitive resource depletion via circadian mismatch does *not* affect choice consistency

In summary, we find that distance to rational behavior across circadian matched and mismatched subjects is similar regardless of the test of rational behavior we conduct. When testing the data's consistency with GARP, for which there are some CCEI benchmarks in the literature, both our circadian matched and mismatched groups would be deemed "rational." We now examine if consistency in behavior, whether cognitively challenged or not, also implies that choices in the risk task are the same.

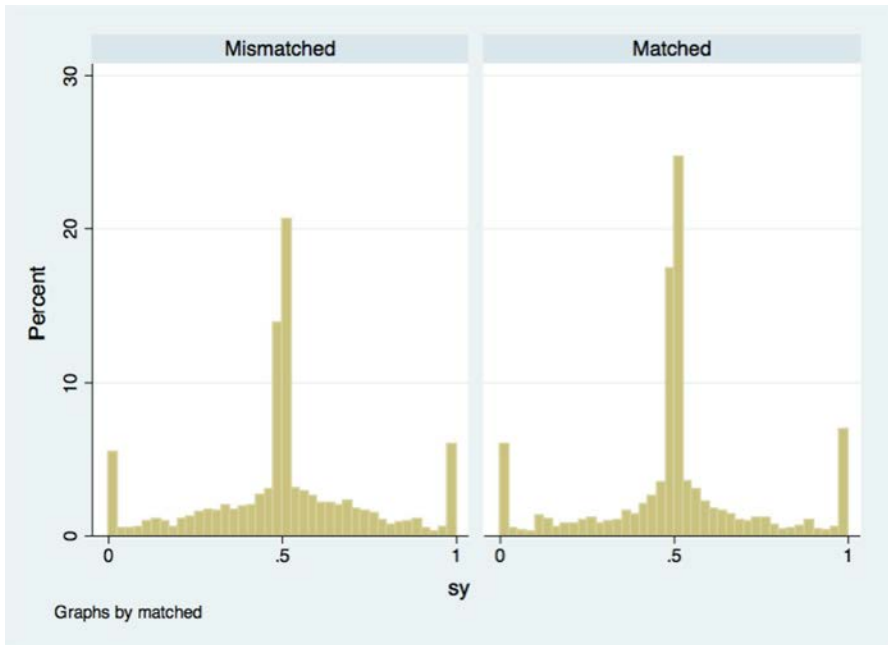


Fig. 3 Share of assets in Y for all relative prices: matched and mismatched subjects

3.3 Choices in the risk task

In light of the absence of a discernible difference in consistency of behavior, in this section, we investigate whether the circadian mismatch manipulation affects risky choices. We look at the distribution of asset investments, from which we calculate certainty equivalents for subjects. These certainty equivalents constitute a theoretically valid measure of risk preference.

Because deviations from expected utility theory might manifest through nonlinear responses to prices, we take these factors into consideration in the analysis that follows. In particular, subjects might choose a distribution of assets that favors constant payoffs. Therefore, small variations in relative prices will have a different impact on asset allocations than large changes in relative prices. It will be important to examine the behavior in these extremes.

Figure 3 shows the distribution of the proportion of the budget share in asset Y for all relative prices for matched and mismatched subjects. Matched subjects tend to more frequently choose asset allocations that secure equal payoffs across states of nature. Thus, matched subjects choose the safe bundle more frequently, which is an indication of increased risk aversion relative to circadian mismatched subjects.¹⁸

¹⁸ Given the mouse-driven graphical choice interface, one might think that sleepy subjects would be less likely to choose safe asset bundles due to motor skill deficits resulting from fatigue. We note that this is not likely the case in our data. However, because that argument would imply that these same sleepy subjects are more likely to choose extreme border asset bundles. This is not the case in our data.

Table 4 Interquartile regression on share in Y ($Y/(Y+X)$)

Variables	No bootstrap	Bootstrap at the subject level
Mismatched	0.085*** [0.012] (0.000)	0.085** [0.042] (0.041)
Constant	0.107*** [0.008] (0.000)	0.107*** [0.023] (0.000)
Observations	10,094	10,094

Standard errors in brackets, p values in parentheses. Only decisions over positive prices are included
 *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

The significance of this result is tested and shown in Table 4. This table shows the results of an interquartile regression of budget share on a dummy variable for being a mismatched subject. The interquartile range of budget share in the Y asset is 8.5 % points larger than that of matched subjects. Mismatched subjects are significantly more likely to make riskier investments.¹⁹

To more rigorously assess differences in risk aversion, we calculate non-parametric certainty equivalents for the matched and mismatched subject groups. The certainty equivalent gives us the amount of money an individual would be indifferent to receiving for sure instead of taking the lottery. Because a relatively more risk averse individual would be willing to take a lower amount of money than the expected value of the lottery, the certainty equivalent gives another measure of risk aversion. To implement this idea, we estimate the certainty equivalent for a particular risky lottery as the highest payoff “safe bundle” (i.e., equal payoffs across states) that would be less preferred to the non-equal payoff “risky” lottery and still be consistent with GARP. That is, we look for all sets of relative prices that would support the “risky” lottery without violating GARP and pick the set of relative prices that include the highest possible “safe bundle.” This highest “safe bundle” is our certainty equivalent measure. So, the more risk averse an individual is, the lower the “safe bundle” would be to make them switch. Finally, calculating certainty equivalents requires subjects to be consistent. So, since many subjects have some violation of GARP, we adjust the revealed preferred relation according to the subject’s CCEI. In particular, all the calculations define a bundle x at prices p to be revealed preferred to y if $CCEI(i) \cdot p \cdot x \geq p \cdot y$, where $CCEI(i)$ is subject i ’s CCEI.²⁰

¹⁹ If we add a dummy variable for being female as an additional control in the regressions in Table 4, our main results still hold. The coefficient on being mismatched is 0.082 (p value = 0.000 in the non-bootstrapped regressions and = 0.062 in the bootstrapped regressions) and the coefficient on female is -0.023 (p value = 0.059 in non-bootstrapped regression and = 0.602 in the bootstrapped regressions). Women are more risk averse than men, but the effect is not always significant. Results are also robust if we control for amount of sleep the night before the experiment or the subject’s CRT score.

²⁰ As a robustness check, we evaluate our results on the differences in certainty equivalents between matched and mismatched subjects by also constraining the sample to subjects whose CCEI is close to one. Our main result, that matched subjects are more risk averse, still holds. The result is no longer statistically significant though, and that is a reflection of the smaller number of observations in the constrained sample.

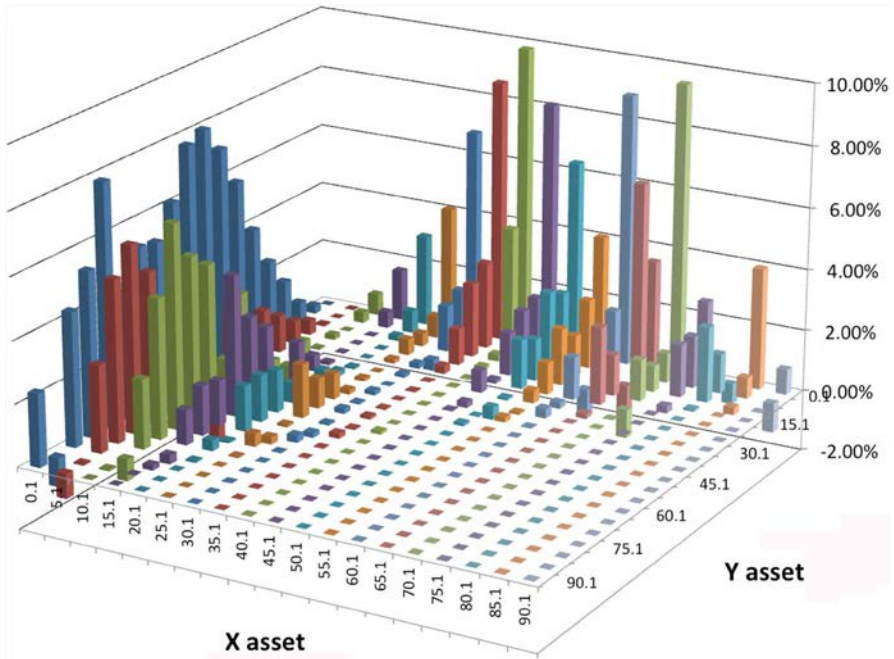


Fig. 4 Differences in certainty equivalents (CE) for different asset bundles as a proportion of the expected value of the lottery for mismatched versus matched subjects

The results from calculating these certainty equivalents are shown in Fig. 4. The average certainty equivalent (CE) for matched and mismatched subjects is calculated for each asset bundle, and Fig. 4 shows the difference of these CEs between mismatched and matched subjects relative to the expected value of the lottery. This shows how much larger the CE is for mismatched compared to matched subjects in percentage points. The general path of where these differences lie is along the range of actual available lotteries seen in the experiment. What is clear from these data is that mismatched individuals have higher calculated certainty equivalents than matched individuals, especially for extreme lotteries that are far from the sure payoff lottery. In other words, matched subjects are more risk averse than mismatched subjects.

We test for the significance of these differences in Table 5. This table shows results for Tobit regressions of the CE on a dummy variable for being mismatched for various sets of relative prices.²¹ The first column includes all the data and shows that mismatched subjects have a 4.4 % point larger CE than matched subjects. The second column confirms this results when the data are restricted to include relative prices strictly different from one. Columns 3–6 further restrict the data to include only progressively

²¹ The Tobit regression takes the calculated certainty equivalent for an individual for a given lottery and regresses it on a dummy variable for being mismatched. There are 361 possible generated lotteries for each individual and 202 individuals, yielding 72,922 observations, and the regressions cluster at the individual level.

Table 5 Tobit estimates of the certainty equivalents as a proportion of the expected value of the lottery

Variables	All lotteries	X I= Y	$\ln(x/y) > \ln(1/1.1) $	$ \ln(x/y) > \ln(1/10) $	$ \ln(x/y) > \ln(1/100) $
Mismatched	0.044*** [0.006] (0.000)	0.044*** [0.006] (0.000)	0.044 [0.033] (0.178)	0.067 [0.049] (0.172)	0.088 [0.061] (0.148)
Constant	3.188*** [0.007] (0.000)	3.186*** [0.007] (0.000)	1.890*** [0.129] (0.000)	2.174*** [0.167] (0.000)	2.188*** [0.149] (0.000)
Observations	72,922	71,710	65,852	10,504	6868

Dummy variables per lottery included, 202 clusters (subjects). Robust standard errors in brackets, p values in parentheses

Table 6 Price elasticity on $y/(y+x)$

Variables	
$\ln(p_y/p_x)$	-0.160*** [0.007] (0.000)
$\ln(p_y/p_x)*\text{Mismatched}$	-0.026** [0.012] (0.028)
Mismatched	0.002 [0.007] (0.710)
Constant	0.503*** [0.005] (0.000)
Observations	10,094
R-squared	0.387

Robust standard errors in brackets, p value in parentheses. Errors clustered at the subject level. Only decisions over positive prices are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

more extreme relative prices. Mismatched subjects have higher CE than matched subjects, but this is only significant in a one-sided test. This leads to our second result:

Result 2 Cognitive resource depletion via circadian mismatch leads to higher certainty equivalents (i.e., increased preference for risk).

As a further check, we examine how price responsiveness varies across groups. Table 6 shows regressions of price elasticity on relative prices, a dummy for being mismatched and an interaction term of relative prices on being mismatched. The significant negative coefficient on the interaction term indicates that mismatched subjects appear more sensitive to changes in relative asset prices than matched subjects. This implies that more steep (flat) budget lines in Fig. 1 lead mismatched subjects to alter their chosen asset bundles even further towards the cheaper asset than would a matched subject.

Another interpretation of this result is that mismatched subjects are less sensitive to the risk of the low (zero) payoff than matched subjects when budget lines reflect large relative price differences. It is also the case that any budget slope other than -1 would lead a risk neutral expected value maximizing subject to choose a corner solution. The Table 6 result is, therefore, consistent with the hypothesis that mismatched subjects take more risks because they are more desensitized to risk, which is a result reported previously in the literature using totally sleep-deprived subjects (McKenna et al. 2007).

In sum, we find that our sleepiness manipulation leads to lower risk aversion as measured by certainty equivalent. This result is consistent with other literature examining extreme forms of temporary cognitive resource depletion effects on incentivized risky choice tasks, such as total sleep deprivation (e.g., McKenna et al. 2007; Ferrara et al. 2015) or intoxication (e.g., Lane et al. 2004). Importantly, despite the shift in risk attitudes, we do not find any significant difference in decision-making rationality resulting from circadian mismatch, under several alternative definitions of consistency. The simultaneous examination of risky choice *and* rationality is an important contribution of our paper, because we document that the increased tendency to take monetary risk due to circadian mismatch is not due to increased randomness of decisions.

4 Conclusions

In this paper, we investigate how a particular form of cognitive resource depletion impacts choice consistency and outcomes in a risky choice task. The task (Choi et al. 2007) generates data that allow us to evaluate choice consistency with respect to several different measures of rationality. As a result, our contribution is that we are able to establish whether differences in preferences over risky asset bundles are the result of “irrationality”, or whether they are the result of state-dependent preferences. We find that they are due to the latter. The circadian mismatch protocol we implement to manipulate cognition is not only effective, but also externally valid and similar to what decision makers face in field environments. While much of the recent literature has focused on how permanent cognitive levels may correlate with risk preferences, we address how temporary fluctuations in available cognitive resources may affect choice, independent of permanent abilities.

Our results are significant and reveal evidence that randomly assigned circadian mismatch subjects are no less rational than matched subjects, and yet preferences for risk shift. Specifically, we have shown that choices are no more or less consistent with GARP or payoff dominance theories as a result of the subject being circadian mismatched. And yet, these mismatched subjects are more willing to accept risky asset bundles compared to matched subjects.

This is an important result with practical and policy implications, especially if one considers that many real-world decision makers face even more serious bouts of sleepiness than the relatively mild manipulation we implement. In the realm of monetary risk choice, sleep deprivation is estimated to affect over 25 % of workers in the financial and insurance industries (CDC 2012). In such industries, any increased tendency to take risk may have significant consequences. In other occupations, risky choice may not involve explicit monetary risk (e.g., air traffic controllers, long-haul trucking, medical practice, or emergency service workers), but sleepiness is commonplace and

of great concern to policymakers establishing regulations that may involve prescribed rest or time-off to avoid sleep deprivation or limit shift work.

If one considers the other various forms of temporary cognitive challenges we often face (e.g., multi-tasking, stress, time pressure), this research may have even more wide reaching implications. We leave it to future research to establish the relationship, if any, between various distinct forms of cognitive resource manipulations, or between cognition effects on monetary risk preference versus other choice domains. Nonetheless, it is clear that this area of research is fertile ground for studying choice in the real world where cognitive resources are not uniformly available at all times.

Acknowledgements The authors thank David Bruner, Olivier l'Haridon, participants at the Economic Science Association meetings, and seminar participants at the University of Rennes and Appalachian State University for helpful comments on earlier drafts of this paper.

Appendix 1: Subject Instruction

Instructions

This is an experiment in decision making. Your payoffs will depend partly on your decisions and partly on chance. Your payoffs will not depend on the decisions of the other participants in the experiment. Please pay careful attention to the instructions as a considerable amount of money is at stake.

The entire experiment should be completed within an hour and a half. At the end of the experiment, you will be paid privately. Your total payoff in this experiment will consist of \$5 as a participation fee (simply for showing up on time), plus whatever payoff you receive from the decision experiment. Details of how your payoff will depend on your decisions will be provided below.

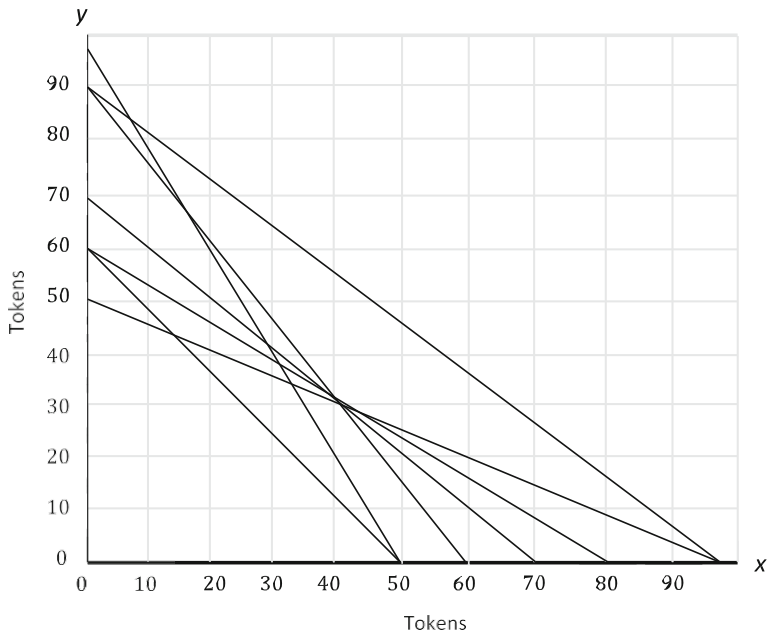
During the experiment, we will speak in terms of experimental tokens instead of dollars. Your payoffs will be calculated in terms of tokens and then translated at the end of the experiment into dollars at the following rate:

$$2 \text{ Tokens} = 1 \text{ Dollar}$$

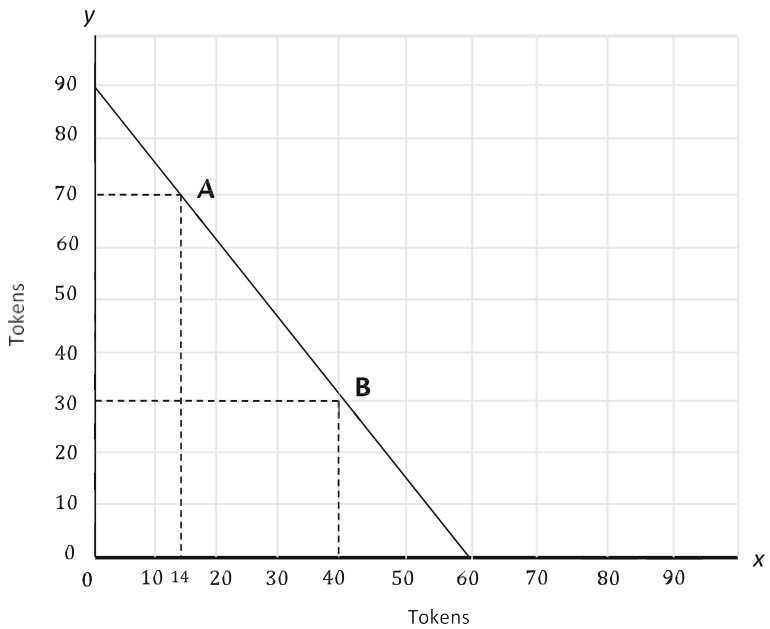
The decision problem

In this experiment, you will participate in 50 independent decision problems that share a common form. This section describes in detail the process that will be repeated in all decision problems and the computer program that you will use to make your decisions.

In each decision problem, you will be asked to allocate tokens between two accounts, labeled x and y . The x account corresponds to the x -axis and the y account corresponds to the y -axis in a two-dimensional graph. Each choice will involve choosing a point on a line representing possible token allocations. Examples of lines that you might face appear in the graph below. Many lines are shown on the same graph to highlight that there will be a variety of different lines you could face, but each decision you make will involve only one line on the graph, as you will see further in these instructions.

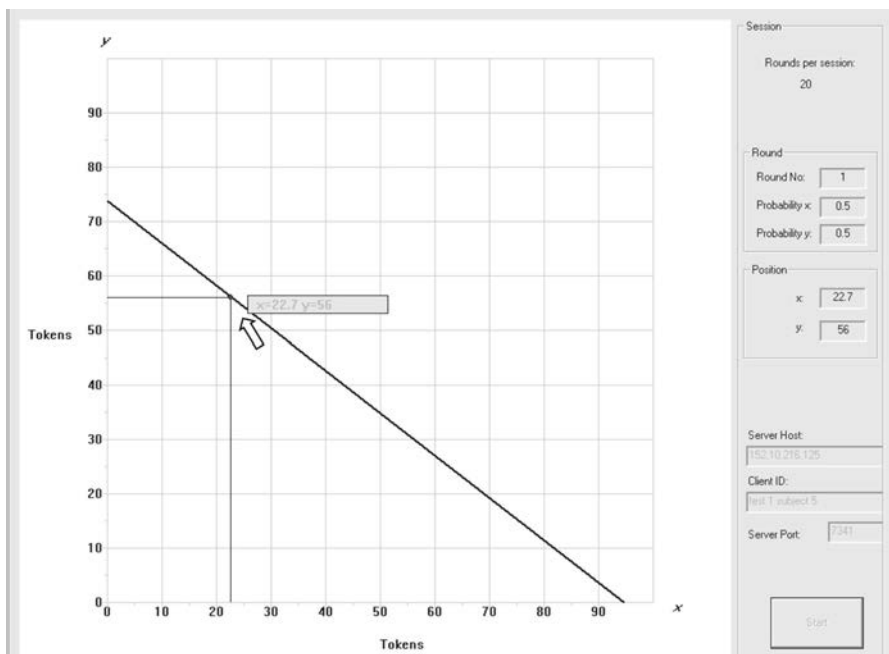


In each choice, you may choose any x and y pair that is on the line. For example, as illustrated in the next graph below, choice **a** represents a decision to allocate 14 tokens in the x account and 70 tokens in the y account. Another possible allocation is **b**, in which you allocate 40 tokens in the x account and 30 tokens in the y account.



Each decision problem will start by having the computer select such a line randomly from the set of lines that intersect with at least one of the axes at 50 or more tokens but with no intercept exceeding 100 tokens. The lines selected for you in different decision problems are independent of each other and independent of the lines selected for any of the other participants in their decision problems.

To choose an allocation in each decision problem, use the mouse to move the pointer on the computer screen to the allocation that you desire. When you are ready to make your decision, left-click to enter your chosen allocation. After that, confirm your decision by clicking on the Submit button that will appear after your decision is made. Note that you can choose only x and y combinations that are on the line (you may also choose either endpoint on any line if you so desire). The next graph shows a picture of the actual decision screen you will see in the experiment. Notice that where you position the pointer on the line will highlight exactly what combination of x and y is at that location on the line. This same information is also shown in the information area to the right of the graph (the example graph shows additional information that will be discussed next in these instructions. It also indicates a 20 round experiment, although today's experiment will be 50 rounds in length).



Once you have confirmed your choice for that decision round, press the OK button. Your payoff in each decision round is determined by the number of tokens in your x account and the number of tokens in your y account. At the end of the round, the computer will randomly select one of the accounts, x or y . There is an equal chance that either account will be selected, and this random selection occurs separately and independently of each participant. You will only receive as payment the number of

tokens you allocated to the account that was chosen. (In the example graph directly above, if account x is selected you would receive 22.7 tokens, and if account y is selected you would receive 56 tokens). The random selection of account x or y in a decision round will not be shown to you until the very end of the experiment.

Once a decision round is finished, you will be asked to make an allocation in another independent decision. This process will be repeated until all 50 decision rounds are completed. At the end of the last round, you will be informed that the experiment has ended.

Your earnings

Your earnings in the experiment are determined as follows. At the end of the experiment, the computer will randomly select one decision round to carry out (that is, 1 out of 50). The round selected depends solely upon chance, and it is equally likely that any round will be chosen. Once a round is chosen, you will receive the number of tokens you allocated to the account (x or y) that was randomly selected for that round. Keep in mind that there is an equal chance that account x or y will be chosen for your token payoff in any given round.

The round selected, your choice and your payment (in terms of tokens) will be shown in the large window that appears at the center of the program dialog window. At the end of the experiment, the tokens will be converted into money. Each token will be worth 0.5 dollars (in other words, your “tokens” payoff will be divided by 2 to get your payoff in dollars). Your final cash earnings in the experiment will be your earnings in the round selected plus the \$5 show-up fee. You will receive your payment as you leave the experiment.

Rules

Please do not share your decisions with anyone else in today’s experiment, please do not talk with anyone during the experiment, and please remain silent until everyone is finished. If there are no further questions, you are ready to start, and an experimenter will start your experiment program.

Appendix 2A: Pre-experiment online survey (answered prior to recruitment)

(some responses elicited using slider bars, matrix response boxes, or other standard online survey features)

What is your gender?

- Female
- Male

Please enter your valid email address (e.g., johndoe@emailserver.com)

This is required if you wish to be entered into the database for eligibility and recruitment consideration for future cash-compensation research experiments

What is your ethnicity?

- Hispanic or Latino
- Not Hispanic or Latino

What is your racial category?

- American Indian/Alaska Native
- Asian
- Native Hawaiian or Other Pacific Islander
- Black or African American
- White (Caucasian)
- Mixed
- Other (please specify in text box)

What is your age?

Are you a student, faculty, or staff?

Over the last 2 weeks, how often have you been bothered by the following problems? (Options for each are “Not at all”, “Several Days”, “Over Half of the Days”, or “Nearly Every Day”)

- Feeling nervous, anxious or on edge
- Not being able to stop or control worrying
- Worrying too much about different things
- Little interest or pleasure in doing things
- Trouble relaxing
- Being so restless that it is hard to sit still
- Becoming easily annoyed or irritable
- Feeling down, depressed, or hopeless
- Feeling afraid as if something awful might happen.

Considering only your own “feeling best” rhythm, at what time would you get up if you were entirely free to plan your day?

- Between 5:00 and 6:30 a.m.
- Between 6:31 and 8:00 a.m.
- Between 8:01 and 9:30 a.m.
- Between 9:31 and 11:00 a.m.
- Between 11:01 a.m. and noon

During the first half-hour after waking up in the morning, how tired do you feel?

- Very tired
- Fairly tired
- Fairly refreshed
- Very refreshed

At what time in the evening do you feel tired and, as a result, in need of sleep?

- Between 8:00 pm and 9:00 pm
- Between 9:01 pm and 10:30 pm
- Between 10:31 pm and 12:30 am
- Between 12:31 am and 2:00 am
- Between 2:01 am and 3:00 am

At what time of the day do you think you reach your “feeling best” peak?

- Between midnight and 4:30 a.m.
- Between 4:31 a.m. and 7:30 a.m.
- Between 7:31 a.m. and 9:30 a.m.
- Between 9:31 a.m. and 4:30 p.m.
- Between 4:31 p.m. and 9:30 p.m.
- Between 9:31 p.m. and midnight.

One hears about “morning” and “evening” types of people. Which ONE of these types do you consider yourself to be?

- Definitely a ‘morning-type’
- More likely a ‘morning-type’ than an ‘evening-type’
- More likely an ‘evening-type’ than a ‘morning-type’
- Definitely an ‘evening-type’

Over the last 7 nights, what is the average amount of sleep you obtained each night?

Last night, how much sleep did you get?

What do you feel is the optimal amount of sleep for you personally to get each night? (optimal in terms of next day alertness, performance, and functionality for you personally.)

How likely are you to doze off or fall asleep in the following situations, in contrast to just feeling tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently, try to work out how they would have affected you.

(options are “would NEVER doze or fall asleep”, “SLIGHT chance of dozing or falling asleep”, “MODERATE chance of dozing or falling asleep”, or “HIGH chance of dozing or falling asleep”)

- Sitting and reading
- Watching TV
- Sitting, inactive in a public place (e.g., a theater or a meeting)
- As a passenger in a car for an hour
- Lying down to rest in the afternoon when circumstances permit
- Sitting and talking to someone
- Sitting quietly after lunch without alcohol
- In a car, while stopped for a few minutes in traffic.

Do you have a diagnosed sleep disorder?

- Yes
- No
- Not diagnosed, but I believe I may have a sleep disorder

If you responded “Yes” (i.e., you have a diagnosed sleep disorder), what is it?

Appendix 2B: Survey given during experiment

APPENDIX 2B-Survey given during experiment

DECISION DAY SURVEYS

SUBJECT NUMBER/CODE _____

Date _____

PLEASE ANSWER BEFORE DECISION TASKS

Please circle the number that best corresponds to how sleepy you feel right now. You may mark any number, but mark only one number (numbers in between those with descriptions correspond to feeling in between the two identified states).

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Extremely Alert		Alert		Neither Alert nor Sleepy		Sleepy		Extremely Sleepy (fighting sleep)

.....*

Date _____



*****PLEASE ANSWER THIS QUESTION AFTER DECISION TASKS*****



For the following three questions, please be as precise as possible (e.g. hours and minutes).

Over the last 7 nights, what is the *average* amount of sleep you obtained each night? _____

Last night, how much sleep did you *get*? _____

What do you feel is the optimal amount of sleep for you personally to get each night? (optimal in terms of next day alertness, performance, and functionality) _____



Please answer the following 3 questions

A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball.
How much does the ball cost?

My response _____

If it takes S machines S minutes to make S widgets, how long would it take
100 machines to make 100 widgets?

My response _____

In a lake, there is a patch of lily pads. Every day, the patch doubles in size.
If it takes 48 days for the patch to cover the entire lake, how long would it
take for the patch to cover half of the lake?

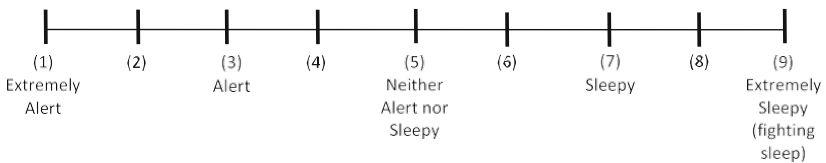
My response _____

How likely are you to doze off or fall asleep in the following situations, in contrast to just feeling tired? **This refers to your usual way of life in recent times.** Even if you have not done some of these things recently, try to work out how they would have affected you.

	Would NEVER doze or fall asleep	SLIGHT chance of dozing or falling asleep	MODERATE chance of dozing or falling asleep	HIGH chance of dozing or falling asleep
Sitting and reading				
Watching TV				
Sitting, inactive in a public place (e.g., a theater or a meeting)				
As a passenger in a car for an hour				
Lying down to rest in the afternoon when circumstances permit				
Sitting and talking to someone				
Sitting quietly after a lunch without alcohol				
In a car, while stopped for a few minutes in traffic				

Please circle the number that best corresponds to how sleepy you feel **right now**. You may mark any number, but mark only one number (numbers in between those with descriptions correspond to feeling in between the two identified states).

***This is the same question you answered at the beginning of today's experiment session. Please respond again. You may or may not choose to give the same response you gave earlier**



Appendix 2C: Technical details of preference function tests

1. Testing expected utility theory.

A test of expected utility in asset markets has been proposed by [Varian \(1983\)](#), and a maintained assumption of Varian's test is that the Bernoulli utility function over money is concave, i.e., individuals are risk averse. Recently, [Polisson et al. \(2015\)](#) have proposed an alternative test of expected utility that does not rely on assuming risk aversion.

The test by Polisson et al. consists of finding a set of numbers, i.e., utilities, that are consistent with the observed revealed preference relation and with an extension of the observed revealed preference relation over the lattice of points generated by the experiment.²² Polisson et al. also propose a way to calculate a number similar to Afriat's critical cost to efficiency corresponding to the test of expected utility theory.

We use the approach by Polisson et al. to test for adherence to expected utility theory.

2. Testing rank-dependent utility.

To test for rank-dependent utility in asset markets, we propose a straightforward generalization of Varian's (1983) test of expected utility to allow for the possibility of rank-dependent valuation of outcomes.

It should be noted that rank-dependent utility can be tested in asset market experiments even if the probability of states of nature does not vary. A simple example will illustrate this point. Consider we observe the following three revealed preference relations: $(x_1, y_1)R(x_2, y_2)$, $(x_3, y_2)R(x_1, y_3)$ and $(x_2, y_3)R(x_3, y_1)$ where aRb means a is revealed preferred to b . We also assume that $x_1 < y_1$, $x_2 > y_2$, $y_1 > x_3$, $x_3 > y_2$ and $x_2 > y_3$.

These allocations are not consistent with expected utility theory. To see this, suppose u rationalizes the choices above when the probability of obtaining asset x is p ($0 < p < 1$). We must have that:

- i. $u(x_1)p + u(y_1)(1-p) > u(x_2)p + u(y_2)(1-p)$
- ii. $u(x_3)p + u(y_2)(1-p) > u(x_1)p + u(y_3)(1-p)$
- iii. $u(x_2)p + u(y_3)(1-p) > u(x_3)p + u(y_1)(1-p)$

Adding inequalities (i) and (ii), we obtain that $u(x_3)p + u(y_1)(1-p) > u(x_2)p + u(y_3)(1-p)$, which contradicts inequality (iii).

However, these inequalities do not contradict rank-dependent utility theory. Let $w(p)$ be the probability weight of p . This theory requires that

- i. $(1-w(1-p))u(x_1) + w(1-p)u(y_1) > (1-w(p))u(y_2) + w(p)u(x_2)$
- ii. $(1-w(p))u(y_2) + w(p)u(x_3) > (1-w(1-p))u(x_1) + w(1-p)u(y_3)$
- iii. $(1-w(p))u(y_3) + w(p)u(y_2) > (1-w(1-p))u(x_3) + w(1-p)u(y_1)$

Consider inequality (i). By assumption, $x_1 < y_1$ and x_1 obtains with probability p and y_1 obtains with probability $1-p$. Rank-dependent utility theory weighs the lower ranked outcome, x_1 , with weight $1-w(1-p)$, with the probability that outcomes are larger than x_1 ($w(1) = 1$) minus the probability that outcomes are strictly larger than x_1 ($w(1-p)$). Similarly, rank-dependent utility theory weighs the largest outcome, y_1 , with the probability that outcomes are larger than y_1 ($w(1-p)$) minus the probability that outcomes are strictly larger than y_1 ($w(0) = 0$). The remaining expressions are obtained in a similar manner.

²² In particular, suppose a person is observed choosing allocation (x,y) in menu 1 and (a,b) in menu 2 for an experiment with two essential Arrow-Debreu securities. The observed quantities in this experiment are $A = \{0,x,y,a,b\}$ and the associated lattice to the experiment is $A \times A$. A preference relation can be extended to this lattice by checking whether x in $A \times A$ was affordable when (x,y) in menu 1 was chosen or when (a,b) in menu 2 was chosen.

Adding up inequalities (i) and (ii), we obtain that $w(1-p)u(y_1) + w(p)u(x_3) > w(1-p)u(y_3) + w(p)u(x_2)$. This inequality would violate rank-dependent utility theory if $w(1-p) = 1-w(p)$ which is not true in general. It is easy to show that if either $x_i > y_i$ or $x_i < y_i$ for $i = 1, 2, 3$, this pattern of behavior would also violate rank-dependent utility.

In sum, it is possible to test for rank-dependent utility theory separated from expected utility even if probabilities are fixed. This follows from the fact that probability weighting depends on the rank of outcomes.

We now show how to test for rank-dependent utility theory in the context of two essential Arrow–Debreu securities. Let the consumption data be $(x_{ik}, p_{ik}, \pi_{ik}), i = 1, 2, k = 1, \dots, N$, where x_{ik}, p_{ik}, π_{ik} are consumption bundles of state contingent assets, their prices and the probability of each state i .

In the context of two state-contingent assets, rank-dependent utility theory is equivalent to the following representation of preferences:

$$\begin{aligned} u(x_{1k}, x_{2k}) &= (1 - w(\pi_{2k}))u(x_{1k}) + w(\pi_{2k})u(x_{2k}) \text{ if } x_{1k} \leq x_{2k} \text{ and} \\ u(x_{1k}, x_{2k}) &= w(\pi_{1k})u(x_{1k}) + (1 - w(\pi_{1k}))u(x_{2k}) \text{ if } x_{1k} > x_{2k} \end{aligned}$$

where u is a monotone increasing value function and w is a monotone increasing probability weighting function with $w(0) = 0$ and $w(1) = 1$.

Because rank-dependent utility theory evaluates lottery prizes according to their ranks, the utility representation of preferences is not differentiable everywhere. In particular, the corresponding indifference curves will have a kink whenever $x_{1k} = x_{2k}$.

To derive the testable hypotheses of rank-dependent utility with a concave value function, we note that if u is differentiable a consumer choosing (x_{1k}, x_{2k}) , where $x_{1k} > x_{2k}$ solves the following problem:

$$\max_{x_{1k}, x_{2k}} \{w(\pi_{1k})u(x_{1k}) + (1 - w(\pi_{1k}))u(x_{2k}) : p_{1k}x_{1k} + p_{2k}x_{2k} \leq 1\}$$

with a symmetric representation for the case in which $x_{1k} < x_{2k}$.

Whenever $x_{1k} = x_{2k}$, the problem is equivalent to the simultaneous solution of the following problems:

$$\begin{aligned} \max_{x_{1k}, x_{2k}} \{w(\pi_{1k})u(x_{1k}) + (1 - w(\pi_{1k}))u(x_{2k}) : p_{1k}x_{1k} + p_{2k}x_{2k} \leq 1, \text{ if } x_{1k} \geq x_{2k}\}, \text{ and} \\ \max_{x_{1k}, x_{2k}} \{(1 - w(\pi_{2k}))u(x_{1k}) + w(\pi_{2k})u(x_{2k}) : p_{1k}x_{1k} + p_{2k}x_{2k} \leq 1, \text{ if } x_{1k} \leq x_{2k}\} \end{aligned}$$

If the function u is concave, first-order conditions are necessary and sufficient for the existence of a maximum. For decisions with $x_{1k} > x_{2k}$, we will have that

$u^t(x_{1k}) = \lambda_k p_{1k} / w(\pi_{1k})$ and $u^t(x_{2k}) = \lambda_k p_{2k} / (1 - w(\pi_{1k}))$, and for decisions with $x_{1k} < x_{2k}$, we will have that $u^t(x_{1k}) = \lambda_k p_{1k} / (1 - w(\pi_{2k}))$ and $u^t(x_{2k}) = \lambda_k p_{2k} / w(\pi_{2k})$. Where λ_k denotes the Lagrange multiplier associated with the problem. Whenever $x_{1k} = x_{2k}$ all these conditions hold simultaneously. In such a case, it is possible for $\lambda_{ik}, i = 1, 2$, to vary for each separate case. Diewert (2012) discusses the conditions under which it is possible to find numbers that satisfy the first-order conditions of the problem.

Below we establish the system of linear inequalities associated with rank-dependent utility theory with a concave value function. Since, function u is concave, we know that

$$u(x_{ik}) \leq u(x_{jm}) + u^t(x_{jm})(x_{ik} - x_{jm})$$

Using the first-order conditions of the associated problem, we have that

$$\begin{aligned} u(x_{ik}) &\leq u(x_{im}) + (\lambda_m p_{jm} / w(\pi_{jm})) (x_{ik} - x_{jm}) \text{ if } x_{jm} > x_{-j,m} \\ u(x_{ik}) &\leq u(x_{im}) + (\lambda_m p_{jm} / (1 - w(\pi_{-j,m}))) (x_{ik} - x_{jm}) \text{ if } x_{jm} < x_{-j,m} \end{aligned}$$

In the case $x_{jm} = x_{-j,m}$, both equations above hold for potentially different values of $\lambda_{j,m}$, $j = 1, 2$.

A concave rank-dependent utility representation of preferences exists if there are positive numbers $\{u_{1k}, u_{2k}, \lambda_{1k}, \lambda_{2k}\}$ ²³ and numbers $\{w_{1k}, w_{2k}\}$ ²⁴ between 0 and 1 satisfying the above equations. The proof of the existence of a rank-dependent utility representation of preferences if the above inequalities are satisfied follows standard arguments.

We follow [Diewert \(2012\)](#) in introducing a new variable S ($S \geq 0$) to be added to all the Afriat inequalities above which equals 0 only if the system of equations has a solution. We also follow [Diewert \(2012\)](#) in restricting λ to be larger than one²⁵ and normalizing income to be equal to one. This *slack* variable can be used to assess the fitness of the data with respect to rank-dependent expected utility. Consider that a subject is an expected utility maximizer. Adding up the inequalities, we find

$$\begin{aligned} \pi_{1m} u(x_{1k}) + \pi_{2m} u(x_{2k}) &\leq \pi_{1m} u(x_{1m}) + \pi_{2m} u(x_{2m}) + \lambda_m (p_{1m} x_{1k} + p_{2m} x_{2k} \\ &\quad - p_{1m} x_{1m} - p_{2m} x_{2m}) \end{aligned}$$

and because income is assumed to be equal to one, we have that

$$\pi_{1m} u(x_{1k}) + \pi_{2m} u(x_{2k}) \leq \pi_{1m} u(x_{1m}) + \pi_{2m} u(x_{2m}) + \lambda_m (p_{1m} x_{1k} + p_{2m} x_{2k} - 1)$$

A violation of expected utility would occur if the utility of allocation (x_{1k}, x_{2k}) is preferred to allocation (x_{1m}, x_{2m}) and it was affordable when (x_{1m}, x_{2m}) was chosen. The equation above would then be violated. Suppose that income is adjusted to level $1 - \varepsilon$ such that allocation (x_{1k}, x_{2k}) is not longer affordable. The minimum value that variable ε takes gives a measure of how much the model departs from the theoretical assumptions. Note that variable ε holds a relationship with the slack variable S since $\max_k \{\lambda_k \varepsilon\}$ will solve the equations above. Since λ is always larger than one, the variable S gives us an upper bound of the value of ε . In the analysis below, we use

²³ $\lambda_{1k} = \lambda_{2k}$ whenever x_{1k} and x_{2k} are different. These numbers are restricted to be positive.

²⁴ $w_{1k} = w_{2k}$ if $p_{1k} = p_{2k} = 1/2$. Only one of the two numbers is necessary if x_{1k} and x_{2k} are different.

²⁵ The normalization is possible because the inequalities are homogeneous of degree one in u and λ .

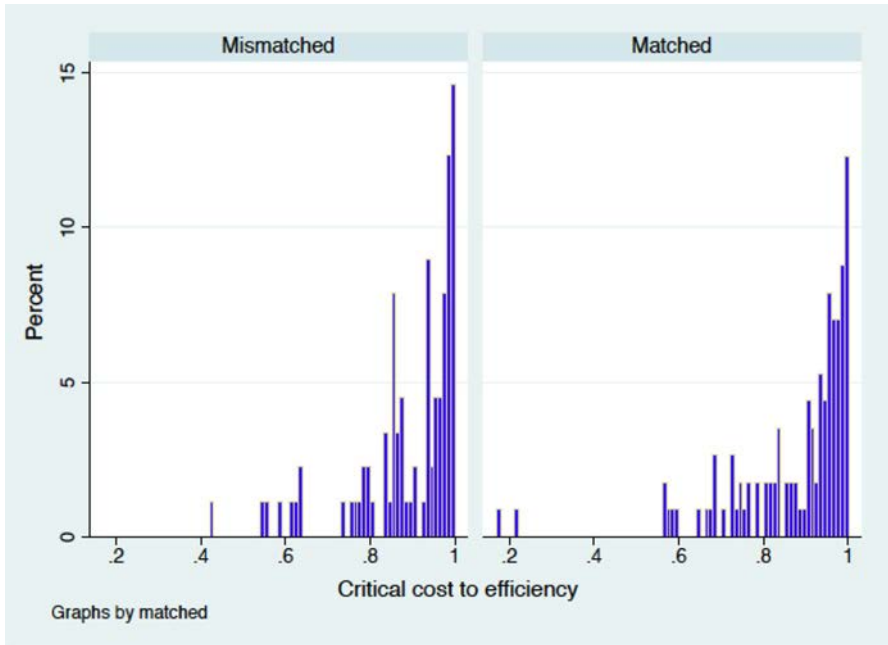


Fig. 5 Distribution of critical cost to efficiency (CCEI) for test of expected utility (Polisson et al. 2015)

1-S as an estimate of the lower bound of the critical cost to efficiency for the test of rank-dependent utility.

3. Results.

Figure 5 shows the distribution of Afriat's critical cost to efficiency (CCEI) corresponding to the test of expected utility developed by Polisson et al. (2015). The mean CCEI of circadian matched subjects is 0.891 and mean CCEI of the circadian mismatch subjects is 0.876. These means are not statistically significantly different (t test = 0.7697, p value = 0.4424, rank sum test p value = 0.5033). Only two subjects, both of them circadian mismatched, have a CCEI equal to 1. This means that only two subjects satisfy conditions for expected utility without any violations of the theory.

Figure 6 shows the distribution of Afriat's critical cost to efficiency (CCEI) corresponding to the test of rank-dependent expected utility with a concave utility function and $w(1/2) \leq 1/2$. While the test presented in Sect. 2 does not require $w(1/2) \leq 1/2$, we consider it to be a natural restriction given the clustering of decisions around the 50-50 allocation. The mean CCEI of circadian matched subjects is 0.797 and mean CCEI of the circadian mismatch subjects is 0.817. These means are not statistically significantly different (t test = 0.7844, p value = 0.4337, rank sum test p value = 0.8434). Only two subjects, both of them circadian mismatched, pass the rank-dependent utility test without any violations (CCEI = 1).

The test of rank-dependent utility implicitly finds a value for a subject's weight for probability 1/2. We find that these estimates are not statistically different according to a t test (0.6869, p value = 0.4930), but they are different according to the rank sum test

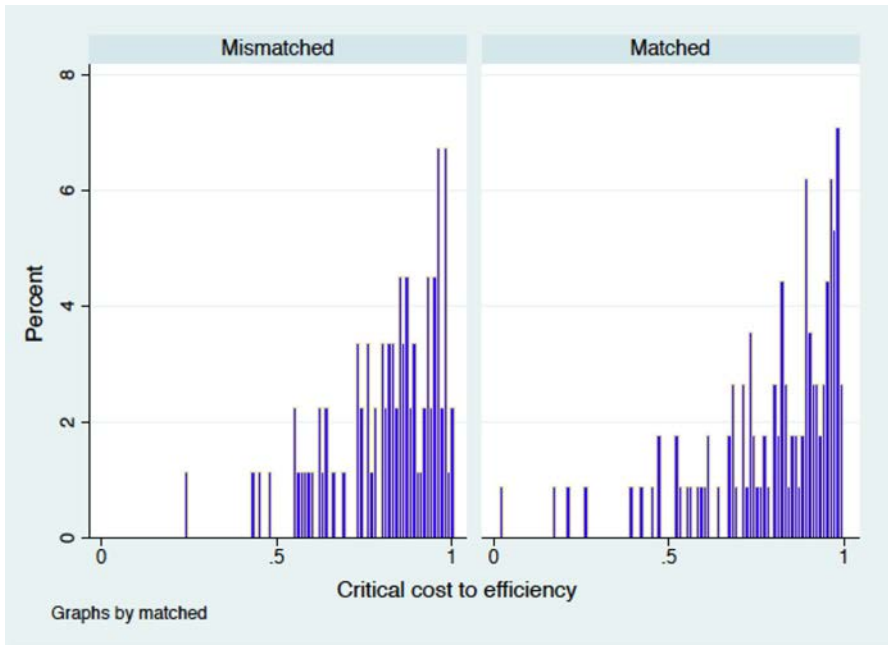


Fig. 6 Distribution of Critical Cost to Efficiency (CCEI) for test of rank-dependent utility with concave utility function and $w(1/2) \leq 1/2$

(p value = 0.0480). To check if there is a difference in estimated probability weight of $1/2$, we test if the proportion having weights below $1/2$ is different across groups. Thirty-three percent of circadian matched subjects have estimated weights below $1/2$ and 21% of circadian mismatched subjects have weights below $1/2$. These proportions are statistically different (t test = -1.9662 , p value = 0.0507). This provides additional support that circadian matched subjects are more likely to choose safe options than circadian mismatched subjects.

References

- Achtziger, A., & Alós-Ferrer, C. (2014). Fast or rational? A response-times study of Bayesian up-dating. *Management Science*, 60, 923–993.
- Adan, A., & Almiral, H. (1991). Horne and Ostberg morningness-eveningness questionnaire: A reduced scale. *Personality and Individual Differences*, 12, 241–253.
- Afriat, S. N. (1972). Efficiency estimation of production function. *International Economic Review*, 8(1), 67–77.
- Becker, G. S., & Murphy, K. M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96(4), 675–700.
- Benjamin, D., Brown, S., & Shapiro, J. (2013). Who is ‘Behavioral’? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, 11(6), 1231–1255.
- Benjamin, D., Choi, J., & Strickland, J. (2010). Social identity and preferences. *American Economic Review*, 100(4), 1913–1928.
- Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, 1, 319–322.

-
- Burghart, D. R., Glimcher, P. W., & Lazzaro, S. C. (2013). An expected utility maximize walks into a bar. *Journal of Risk and Uncertainty*, 46(3), 215–246.
- Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior and job attachment. *Proceeding of the National Academy of Sciences*, 106(19), 7745–7750.
- Callen, M., Isaqzadeh, M., Long, J., & Sprenger, C. (2014). Violence and Risk Preferences: Artefactual and Experimental Evidence from Afghanistan. *American Economic Review*, 104(1), 123–148.
- Centers for Disease Control and Prevention. (2012). Morbidity and Mortality Weekly Report. April 27, 61(16): 281–285. Accessed at <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6116a2.htm> September 27, 2012.
- Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Genetic variation in preferences for giving and risk taking. *Quarterly Journal of Economics*, 124(2), 809–842.
- Chelminski, I., Petros, T. V., Plaud, J. J., & Ferraro, F. R. (2000). Psychometric properties of the reduced Horne and Ostberg questionnaire. *Personality and Individual Differences*, 29(3), 469–478.
- Chelminski, I., Ferraro, F. R., Petros, T., & Plaud, J. J. (1997). Horne and Ostberg Questionnaire: A score distribution in a large sample of young adults. *Personality and Individual Differences*, 23(4), 647–652.
- Choi, S., Fisman, R., Gale, D., & Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, 27(5), 1921–1938.
- Coren, S. (1996). Daylight savings time and traffic accidents. *New England Journal of Medicine*, 334, 924.
- Dickinson, D. L., & McElroy, T. (2010). Rationality around the clock: Sleep and time-of-day effects on guessing game responses. *Economic Letters*, 108, 245–248.
- Dickinson, D. L., & McElroy, T. (2012). Circadian effects on strategic reasoning. *Experimental Economics*, 15(3), 444–459. doi:10.1007/s10683-011-9307-3.
- Diewert, E. (2012). Afriat's theorem and some extensions to choice under uncertainty. *The Economic Journal*, 122(60), 305–331.
- Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(June), 1238–1260.
- Ferrara, M., Bottasso, A., Tempesta, D., Carrieri, M., De Gennaro, L., & Ponti, G. (2015). Gender differences in sleep deprivation effects on risk and inequality aversion: Evidence from an economic experiment. *PLoS One*, 10(3), e0120029.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19, 25–42.
- Garbarino, E., Slonim, R., & Sydnor, J. (2011). Digit ratios (2D:4D) as predictors of risk decision making for both sexes. *Journal of Risk and Uncertainty*, 42(1), 1–26.
- Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, 4, 97–110.
- Koszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4), 1047–1073.
- Krajbich, I., Bartling, B., Hare, T., & Fehr, E. (2015). eThinking fast and slow based on a critique of reaction-time reverse inference. *Nature Communications*, 6, 7455.
- Kruglanski, A. W., & Pierro, A. (2008). Night and day, you are the one. On circadian mismatches and the transference effect in social perception. *Psychological Science*, 19(3), 296–301.
- Lane, S. D., Cherek, D. R., Pietras, C. J., & Tcheremissine, O. V. (2004). Alcohol effects on human risk taking. *Psychopharmacology*, 172, 68–77.
- Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*, 126(1), 373–416.
- Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. New York: Oxford University Press.
- McKenna, B. S., Dickinson, D. L., Orff, H. J., Drummond, S., & Sean, P. A. (2007). The effects of one night of sleep deprivation on known-risk and ambiguous-risk decisions. *Journal of Sleep Research*, 16, 245–252.
- Paine, S. J., Gander, P. H., & Travier, N. (2006). The epidemiology of morningness/eveningness: Influence of age, gender, ethnicity, and socioeconomic factors in adults (30–49 years). *Journal of Biological Rhythms*, 21(1), 68–76.
- Polisson, D., Quah, J. K. H., Renou, L. (2015). *Revealed Preferences over Risk and Uncertainty*. WP 740, Economics Department, Oxford University.

-
- Read, D., Loewenstein, G., & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, 19(1–3), 171–197.
- Smith, C. S., Folkard, S., Schmieder, R. A., Parra, L. F., Spelten, E., Almira, H., et al. (2002). Investigation of morning-evening orientation in six countries using the preferences scale. *Personality and Individual Differences*, 32, 949–968.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- Varian, H. R. (1983). Nonparametric tests of models of investor behavior. *Journal of Financial and Quantitative Analysis*, 18(3), 269–278.
- Venkatraman, V., Chuah, Y. M. L., & Huettel, S. A. (2007). Sleep deprivation elevates expectation of gains and attenuates response to losses following risky decisions. *Sleep*, 30, 603–609.
- Voors, M. J., Nillesen, E. E. M., Verwimp, P., Bulte, E. H., & Lensink, R. (2012). Violent conflict and behavior: A field experiment in burundi. *American Economic Review*, 102(2), 941–964.
- Wozniak, D., Harbaugh, W. T., & Mayr, U. (2014). The menstrual cycle and performance feedback alter gender differences in competitive choices. *Journal of Labor Economics*, 32(1), 161–198.