RANDOM DISCOUNTED EXPECTED UTILITY: ONLINE APPENDIX

JOSE APESTEGUIA†, MIGUEL A. BALLESTER‡, AND ÁNGELO GUTIÉRREZ-DAZA§

Contents

Date: July, 2024.

Appendix A. Proofs

Proof of Proposition 1: Consider a menu $A_{\mathcal{R}} = \{0_{\mathcal{R}}, 1_{\mathcal{R}}\} = \{([p, 1-p; x_1^0, x_2^0], 0), ([p, 1-p; x_1^1, x_2^1], 0)\}$, such that $x_1^1 > x_1^0 > x_2^0 > x_2^1$ and $p \in (0, 1)$. Consider $r < r'$. Construct the affine transformations $v_r, v_{r'}$ of $u_r, u_{r'}$ satisfying $v_r(\omega + x_2^1) = v_{r'}(\omega + x_2^1) = 0$ and $v_r(\omega+x_2^0) = v_{r'}(\omega+x_2^0) = 1$. By strict monotonicity of the original utility functions, it must be $v_r(\omega + x_1^0) > 1$ and $v_{r'}(\omega + x_1^0) > 1$. We now claim that $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$ must hold, and this will be proved by contradiction. Assume that $v_{r'}(\omega + x_1^0) \geq v_r(\omega + x_1^0) > 1$. In this case, we can consider the lotteries $[p^*, 1 - p^*; x_1^0, x_2^1]$ and $[1; x_2^0]$, with $\frac{1}{v_{r'}(\omega+x_1^0)} \leq p^* \leq \frac{1}{v_r(\omega+x_1^0)}$. It is immediate to see that the expected utility constructed upon v_r leads to, at least, weakly prefer lottery $[1; x_2^0]$ while the expected utility constructed upon $v_{r'}$ leads to, at least, weakly prefer lottery $[p^*, 1-p^*; x_1^0, x_2^1]$. This contradicts the fact that $v_{r'}$, being a strict concave transformation of v_r, must have a strictly lower certainty equivalent for the second, riskier lottery and hence, we have proved that $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$.

We now claim that $v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^1) - v_r(\omega + x_1^0)$ must hold, and prove it by contradiction. If it were not true, given that we already proved $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$, we would have $\frac{v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0)}{v_{r'}(\omega+x_1^0)} > \frac{v_r(\omega+x_1^1)-v_r(\omega+x_1^0)}{v_r(\omega+x_1^0)}$. Considering the lotteries $[p', 1-p'; x_1^1, x_2^1]$ and $[1; x_1^0]$, with $\frac{v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0)}{v_{r'}(\omega+x_1^0)} > \frac{1-p'}{p'} > \frac{v_r(\omega+x_1^1)-v_r(\omega+x_1^0)}{v_r(\omega+x_1^0)}$, the expected utility constructed upon v_r would lead to the choice of $[1; x_1^0]$ while the expected utility constructed upon $v_{r'}$ would lead to the choice of $[p', 1 - p'; x_1^1, x_2^1]$. This is again a contradiction with the concavity assumption, and concludes the argument; it must then be $v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^1) - v_r(\omega + x_1^0)$.

We now claim that for every (r, δ) and (r', δ') such that $r' > r$, if $DEU_{r,\delta}(0_{\mathcal{R}}) \geq DEU_{r,\delta}(1_{\mathcal{R}})$, then $DEU_{r',\delta'}(0_{\mathcal{R}}) > DEU_{r',\delta'}(1_{\mathcal{R}})$. To see this, suppose that $DEU_{r,\delta}(0_{\mathcal{R}}) \geq DEU_{r,\delta}(1_{\mathcal{R}})$. This is equivalent to claim that the expected utility of lottery $[p, 1-p; x_1^0, x_2^0]$ is greater than the expected utility of lottery $[p, 1-p; x_1^1, x_2^1]$ when the monetary utility u_r is used. That is equivalent to claim that the expected utility of lottery $[p, 1 - p; x_1^0, x_2^0]$ is greater than the expected utility of lottery $[p, 1-p; x_1^1, x_2^1]$ when the monetary utility v_r is used, and can be written as $\frac{1-p}{p} \geq \frac{v_r(\omega+x_1^1)-v_r(\omega+x_1^0)}{v_r(\omega+x_2^0)-v_r(\omega+x_2^1)} = v_r(\omega+x_1^1)-v_r(\omega+x_1^0)$. From our previous claims, we know that it must be $\frac{1-p}{p} > v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) = \frac{v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0)}{v_{r'}(\omega+x_2^0)-v_{r'}(\omega+x_2^1)}$, which implies that the first lottery is strictly preferred to the second using $v_{r'}$ or, alternatively, using $u_{r'}$. This implies $DEU_{r',\delta'}(0_{\mathcal{R}}) > DEU_{r',\delta'}(1_{\mathcal{R}})$ and concludes the argument. With the unbounded curvature assumption, the certainty equivalent of both lotteries must converge to the maximum and minimum payout when r tends to $-\infty$ and $+\infty$, respectively. That is, there are values of r for which $0_{\mathcal{R}}$ and $1_{\mathcal{R}}$ are preferred. As a result, there must be a unique $K(A_{\mathcal{R}}) \in \mathbb{R}$ such that alternative $0_{\mathcal{R}}$ is preferred if and only if $r \geq K(A_{\mathcal{R}})$ which leads to Claim 1. For Claim 2, note that whenever F^r dominates G^r, it must be $\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) = 1 - F^r(K(A_{\mathcal{R}})) > 1 - G^r(K(A_{\mathcal{R}})) = \mathcal{P}_g(0_{\mathcal{R}}, A_{\mathcal{R}})$. For Claim 3, notice that the assumption requires us to consider two cases, $\text{med}(F^r) > K(A_{\mathcal{R}})$ or $\text{med}(F^r) < K(A_{\mathcal{R}})$. In the first case, since F^r expands G^r, it must be $F^r(K(A_{\mathcal{R}})) > G^r(K(A_{\mathcal{R}})) > 1/2$, while in the second case, it must be that $F^r(K(A_{\mathcal{R}})) < G^r(K(A_{\mathcal{R}})) < 1/2$, concluding the proof. ■

Proof of Proposition 2: Consider a menu $A_{\mathcal{T}} = \{0_{\mathcal{T}}, 1_{\mathcal{T}}\} = \{([1; x^0], t^0), ([1; x^1], t^1)\}$ such that $t^0 < t^1$ and $x^0 < x^1$. From the definition of DEU, it follows immediately that

$$DEU_{r,\delta}(0_{\mathcal{T}}) \geq DEU_{r,\delta}(1_{\mathcal{T}}) \Leftrightarrow \delta \geq K(A_{\mathcal{T}}|r) = \frac{1}{t^1 - t^0} \log \frac{u_r(\omega + x^1)}{u_r(\omega + x^0)}.$$

Strict monotonicity of u_r guarantees that this threshold is always a positive real value and, hence, the first expression in Claim 1, and Claim 2a, follow immediately.

We now claim that the threshold map $\{K(A_{\mathcal{T}}|r)\}_{r \in \mathbb{R}}$ is strictly decreasing in r. To see this, notice that scalar transformations leave DEU decisions unaffected. Hence, we can select the scalar transformation v_r of u_r for which $v_r(\omega + x^1) = 1$ holds and, then, we are only required to show that $\frac{-\log v_r(\omega + x^0)}{t^1 - t^0}$ is strictly decreasing in r or, equivalently, that $v_r(\omega + x^0)$ is strictly increasing in r. Suppose by contradiction that this is not the case, i.e., $v_r(\omega + x^0) \geq v_{r'}(\omega + x^0)$ with $r < r'$. By considering the lotteries $[p^*, 1 - p^*; \omega, \omega + x^1]$ and $[1; \omega + x^0]$, where $v_r(\omega + x^0) \geq p^* \geq v_{r'}(\omega + x^0)$, it is immediate to see that the expected utility, using v_r, is larger for the first lottery than for the second, while the expected utility, using $v_{r'}$, is larger for the second lottery than for the first, a contradiction with the strict concavity assumption. The threshold map is thus strictly decreasing in r.

The unbounded curvature assumption also proves that the map is onto for $\mathbb{R}_{++}$ and hence, it is a bijective map between $\mathbb{R}$ and $\mathbb{R}_{++}$. Thus, it can be inverted to obtain the strictly decreasing thresholds $\{K(A_{\mathcal{T}}|\delta)\}_{\delta\in\mathbb{R}_{++}}$, such that, for a given $\delta>0$, alternative $0_{\mathcal{T}}$ is chosen if and only if r is above this threshold. For $\delta\leq 0$ alternative $1_{\mathcal{T}}$ is always chosen. Hence, $\mathcal{P}_f(0_{\mathcal{T}},A_{\mathcal{T}})=\int_r(1-F_{\delta|r}(K(A_{\mathcal{T}}|r)))f^r(r)\mathrm{d}r=\int_{\delta>0}(1-F_{r|\delta}(K(A_{\mathcal{T}}|\delta)))f^\delta(\delta)\mathrm{d}\delta$, where f^δ is the marginal density of δ. The second expression in Claim 1, and Claim 2b, follow.

For Claim 3a, we just need to reproduce the logic of Proposition 1, expanding separately each of the conditional distributions $F_{\delta|r}$. This always creates a strictly larger conditional stochasticity. From there, we need to prove that the argument extends to the weighted aggregation of all these conditional distributions. To see this, notice that the continuity of the map $\{K(A_{\mathcal{T}}|r)\}_{r\in\mathbb{R}}$ guarantees that all conditional medians of δ lie on the same side of the threshold map. As a result, the same alternative, either $0_{\mathcal{T}}$ or $1_{\mathcal{T}}$, is chosen more often in each of the conditionals, and the expansion argument extends to the aggregation. For Claim 3b, a similar argument holds by expanding the conditionals $F_{r|\delta}$ and using the continuity of the inverse map. ■

Proof of Proposition 3: We start by showing that $\theta_r\equiv(\mu_r,\sigma_r)$ is identified. Assume, on the contrary, that this is not the case: there exist θ'_r and θ^*_r in Θ_r such that $\theta'_r\neq\theta^*_r$ and $\mathcal{P}_{\theta'}(1_{\mathcal{R}},A_{\mathcal{R}})=\mathcal{P}_{\theta^*}(1_{\mathcal{R}},A_{\mathcal{R}})$. Using Proposition 1, we can write this equality as:

$$\mathcal{P}_{\theta'_r}\left(r\leq K\left(A_{\mathcal{R}}\right)\right)=\Phi\left(\frac{K\left(A_{\mathcal{R}}\right)-\mu'_r}{\sigma'_r}\right)=\Phi\left(\frac{K\left(A_{\mathcal{R}}\right)-\mu^*_r}{\sigma^*_r}\right)=\mathcal{P}_{\theta^*_r}\left(r\leq K\left(A_{\mathcal{R}}\right)\right).$$

Since the $\Phi(\cdot)$ is a strictly monotonic function, the last equality implies that:

$$-\frac{\mu^*_r}{\sigma^*_r}+\frac{1}{\sigma^*_r}K\left(A_{\mathcal{R}}\right)=-\frac{\mu'_r}{\sigma'_r}+\frac{1}{\sigma'_r}K\left(A_{\mathcal{R}}\right), \tag{A.1}$$

for every menu $A_{\mathcal{R}}$. Now, since $\theta'_r\neq\theta^*_r$, there are three possible cases: $\mu'_r\neq\mu^*_r$ and $\sigma'_r=\sigma^*_r$; $\mu'_r=\mu^*_t$ and $\sigma'_r\neq\sigma^*_r$; and $\mu'_r\neq\mu^*_r$ and $\sigma'_r\neq\sigma^*_r$. In the first case, equation (A.1) implies $\mu'_r=\mu^*_r$, leading to a contradiction. Consider now the second and third cases where $\sigma'_r\neq\sigma^*_r$. Evaluating (A.1) for $A_{\mathcal{R},a}$ and $A_{\mathcal{R},b}$ and combining the resulting expressions yields:

$$\left(\frac{1}{\sigma_r^*}-\frac{1}{\sigma_r'}\right)\left(K\left(A_b^{\mathcal{R}}\right)-K\left(A_a^{\mathcal{R}}\right)\right)=0. \tag{A.2}$$

Since $K\left(A_{\mathcal{R},a}\right)\neq K\left(A_{\mathcal{R},b}\right)$ by assumption (a), it must be the case that $\sigma_r'=\sigma_r^*$, arriving to a contradiction. We thus conclude that θ_r is identified. The next step is to show that $\theta_\delta\equiv(\mu_\delta,\sigma_\delta,\rho)$ is identified. Fix (μ_r,σ_r) and assume, on the contrary, that θ_δ is not identified: there exist θ_δ' and θ_δ^* in Θ such that $\theta_\delta'\neq\theta_\delta^*$ and $\mathcal{P}_{\theta_\delta'}\left(0_{\mathcal{T}},A_{\mathcal{T}}\right)=\mathcal{P}_{\theta_\delta^*}\left(0_{\mathcal{T}},A_{\mathcal{T}}\right)$. Using Proposition 2, the equality of probabilities implies:

$$\int_{-\infty}^{\infty}\left\{\Phi\left(\frac{K\left(A_{\mathcal{T}}|r\right)-\mu_{\delta|r}'}{\sigma_{\delta|r}'}\right)-\Phi\left(\frac{K\left(A_{\mathcal{T}}|r\right)-\mu_{\delta|r}^*}{\sigma_{\delta|r}^*}\right)\right\}\phi\left(\frac{r-\mu_r}{\sigma_r}\right)dr=0,$$

with $\mu_{\delta|r}=\mu_\delta+\rho\sigma_\delta v\left(r\right)$, $\sigma_{\delta|r}=\sigma_\delta\sqrt{1-\rho^2}$, and $\nu\left(r\right)\equiv\frac{r-\mu_r}{\sigma_r}$. The term in brackets in the previous expression is bounded in $[-1,1]$. By the continuity and monotonicity of $\phi\left(\cdot\right)$ and $\Phi\left(\cdot\right)$, there exists $\overline{r}_m\in\mathbb{R}$ for each one of the menus $\{A_{\mathcal{T},c},A_{\mathcal{T},d},A_{\mathcal{T},e}\}$ such that:

$$-\frac{\mu_\delta'}{\sigma_{\delta|r}'}-\alpha'v\left(\overline{r}_m\right)+\frac{1}{\sigma_{\delta|r}'}K\left(A_{\mathcal{T},m}|\overline{r}_m\right)=-\frac{\mu_\delta^*}{\sigma_{\delta|r}^*}-\alpha^*v\left(\overline{r}_m\right)+\frac{1}{\sigma_{\delta|r}^*}K\left(A_{\mathcal{T},m}|\overline{r}_m\right), \tag{A.3}$$

where $\alpha\equiv\rho/\sqrt{1-\rho^2}$ and we use the fact that $\sigma_{\delta|r}$ is independent of $\overline{r}_m$. Now, any of the three conditions in Assumption (b) implies $K\left(A_{\mathcal{T},c}|r\right)\neq K\left(A_{\mathcal{T},d}|r\right)\neq K\left(A_{\mathcal{T},e}|r\right)\neq K\left(A_{\mathcal{T},c}|r\right)$ for any $r\in\mathbb{R}$. Using this result and the Implicit Function Theorem, we conclude that that $\overline{r}_c\neq\overline{r}_d\neq\overline{r}_e\neq\overline{r}_c$. Setting $m=c$ in (A.3) and subtracting the corresponding expression for $m=d$, we get:

$$\left[\alpha'-\alpha^*\right]\left(v\left(\overline{r}_c\right)-v\left(\overline{r}_d\right)\right)+\left[\frac{1}{\sigma_{\delta|r}'}-\frac{1}{\sigma_{\delta|r}^*}\right]\left(K\left(A_{\mathcal{T},c}|\overline{r}_c\right)-K\left(A_{\mathcal{T},d}|\overline{r}_d\right)\right)=0. \tag{A.4}$$

Repeating this procedure for menus $A_{\mathcal{T},c}$ and $A_{\mathcal{T},e}$, we get:

$$\left[\alpha'-\alpha^*\right]\left(v\left(\overline{r}_c\right)-v\left(\overline{r}_e\right)\right)+\left[\frac{1}{\sigma_{\delta|r}'}-\frac{1}{\sigma_{\delta|r}^*}\right]\left(K\left(A_{\mathcal{T},c}|\overline{r}_c\right)-K\left(A_{\mathcal{T},e}|\overline{r}_e\right)\right)=0. \tag{A.5}$$

Using (A.5) to replace $\left[1/\sigma_{\delta|r}'-1/\sigma_{\delta|r}^*\right]$ in (A.4), we get:

$$\left[\alpha'-\alpha^*\right]\left[\overline{r}_c-\overline{r}_d\right]\left[1-\left(\frac{\overline{r}_c-\overline{r}_e}{\overline{r}_c-\overline{r}_d}\right)\left(\frac{K\left(A_{\mathcal{T},c}|\overline{r}_c\right)-K\left(A_{\mathcal{T},d}|\overline{r}_d\right)}{K\left(A_{\mathcal{T},c}|\overline{r}_c\right)-K\left(A_{\mathcal{T},e}|\overline{r}_e\right)}\right)\right]=0. \tag{A.6}$$

Since $K(A_{\mathcal{T},c}|\overline{r}_c) \neq K(A_{\mathcal{T},d}|\overline{r}_d) \neq K(A_{\mathcal{T},e}|\overline{r}_e) \neq K(A_{\mathcal{T},c}|\overline{r}_c)$, equation (A.6) implies that $\alpha' = \alpha^*$, which in turn implies $\rho' = \rho^*$. This result and equation (A.5) implies that $\sigma'_{\delta|r} = \sigma^*_{\delta|r}$, so that $\sigma'_\delta = \sigma^*_\delta$. Using equation (A.3), we conclude that $\mu'_\delta = \mu^*_\delta$, arriving to a contradiction. This concludes the proof. ■

Proof of Proposition 4: To simplify the presentation of the result and provide neat intuitions, we use two relatively mild assumptions: (i) the probability of the event $\{r > 1, \delta < 0\}$ is small and (ii) there are two time menus $A_{\mathcal{T}_1}$ and $A_{\mathcal{T}_2}$ with the same payouts where the probabilities of selecting options $0_{\mathcal{T}_1}$ and $1_{\mathcal{T}_2}$ are greater than $\frac{1}{2}$. To motivate (i), notice that $\delta < 0$ already corresponds to the rare event in which the individual has a strict preference for the future, and we are compounding this with the extra effect of a more-than-logarithmic curvature. To motivate (ii), notice that we can make the differences in the timings as small or as large as desired.

Now, from Proposition 1 we know that $\mathcal{P}_f(0_\mathcal{R}, A_\mathcal{R}) = 1 - F^r(K(A_\mathcal{R}))$, which in the parametric version reads as $\mathcal{P}_f(0_\mathcal{R}, A_\mathcal{R}) = 1 - \Phi(\frac{K(A_\mathcal{R})-\mu_r}{\sigma_r})$. Fix any value of $p \in (0, 1)$, three payouts $x^0_1 > x^0_2 > x^1_2 > 0$, and consider risk menus that vary only on the payout x^1_1. The thresholds of these menus are strictly increasing in x^1_1 and form a bijection with the real numbers. Hence, there exist two menus $A_{\mathcal{R}_1}$ and $A_{\mathcal{R}_2}$ such that the choice probabilities for options $0_{\mathcal{R}_1}$ and $0_{\mathcal{R}_2}$ are equal to $1 - \Phi(0)$ and $1 - \Phi(1)$, respectively. It must then be $\mu_r = K(A_{\mathcal{R}_1})$ and $\sigma_r = K(A_{\mathcal{R}_2}) - \mu_r = K(A_{\mathcal{R}_2}) - K(A_{\mathcal{R}_1})$.

We now discuss time menus and, by using sufficiently large payouts in all our arguments, we can assume w.l.o.g. that behavior corresponds to the case $\omega \to 0$. In this limit case the conditional threshold map becomes piece-wise linear: (a) when $r < 1$, $K(A_\mathcal{T}|r) = \frac{\log \frac{x^1}{x^0}}{t^1 - t^0}(1 - r) \equiv K(A_\mathcal{T})(1 - r) > 0$, where $K(A_\mathcal{T})$ is a menu-dependent constant, and (b) when $r \geq 1$, $K(A_\mathcal{T}|r)$ becomes null. Hence, the probability of choosing $0_\mathcal{T}$ corresponds to the probability that δ lies above $\min\{0, K(A_\mathcal{T})(1 - r)\}$ and, given our assumption (i), this can be approximated by the probability that δ lies above $K(A_\mathcal{T})(1 - r)$. This is the probability that the random variable $K(A_\mathcal{T})(1 - r) - \delta$ lies below zero. Given normality this random variable is also normal with mean $-\mu_\delta - K(A_\mathcal{T})\mu_r + K(A_\mathcal{T})$ and standard deviation $\sqrt{K^2(A_{\mathcal{T}})\sigma_r^2+\sigma_\delta^2+2\rho K(A_{\mathcal{T}})\sigma_r\sigma_\delta}$. Thus the choice probability of $0_{\mathcal{T}}$ is approximately $\Phi\big(\frac{\mu_\delta+K(A_{\mathcal{T}})\mu_r-K(A_{\mathcal{T}})}{\sqrt{K^2(A_{\mathcal{T}})\sigma_r^2+\sigma_\delta^2+2\rho K(A_{\mathcal{T}})\sigma_r\sigma_\delta}}\big)$.

By assumption (ii) there exist two time menus $A_{\mathcal{T}_1}$ and $A_{\mathcal{T}_2}$ with the same payouts x^0, x^1 such that the choice probabilities of $0_{\mathcal{T}}$ are above and below $\Phi(0)=\frac{1}{2}$, respectively. Due to the stationarity of DEU, we can assume w.l.o.g. that the earlier payout takes place in the present in both cases. By continuity there must exist a unique t, and hence a menu $A_{\mathcal{T}_3}=\{([1;x^0],0),([1;x^1],t)\}$ such that the choice probability of option $0_{\mathcal{T}_3}$ is exactly $\Phi(0)$. Hence, $\mu_\delta=K(A_{\mathcal{T}_3})(1-\mu_r)=K(A_{\mathcal{T}_3})(1-K(A_{\mathcal{R}_1}))$. Second, consider any sequence of time problems $\{A_{\mathcal{T}^n}\}$ such that $\lim_n K(A_{\mathcal{T}^n})=0$. Denote by q the limit of the choice probabilities of option $0_{\mathcal{T}^n}$. We know that $q=\Phi(\frac{\mu_\delta}{\sigma_\delta})$ and, hence, it must be $\sigma_\delta=\frac{\mu_\delta}{\Phi^{-1}(q)}=\frac{K(A_{\mathcal{T}_3})(1-K(A_{\mathcal{R}_1}))}{\Phi^{-1}(q)}$. Finally, by fixing again any three parameters in a time menu and varying the fourth, we know that there exists a unique time menu $A_{\mathcal{T}_4}$ in such a family for which $K(A_{\mathcal{T}_4})=\frac{\sigma_\delta}{\sigma_r}$.[1] Denote by q' the choice probability of option $0_{\mathcal{T}_4}$. It must then be $q'=\Phi(\frac{\frac{\mu_r-1}{\sigma_r}+\frac{\mu_\delta}{\sigma_\delta}}{\sqrt{2(1+\rho)}})$. Notice that the right-hand side map is either strictly increasing or strictly decreasing in ρ, which allows to obtain ρ: $\rho=\frac{1}{2}\Big[\frac{\frac{\mu_r-1}{\sigma_r}+\frac{\mu_\delta}{\sigma_\delta}}{\Phi^{-1}(q')}\Big]^2-1=\frac{1}{2}\Big[\frac{\frac{K(A_{\mathcal{R}_1})-1}{K(A_{\mathcal{R}_2})-K(A_{\mathcal{R}_1})}+\Phi^{-1}(q)}{\Phi^{-1}(q')}\Big]^2-1$. This concludes the proof. ■

Proof of Proposition 5: Consider a menu $A_{\mathcal{C}}$ defined by $(p^0,x^0,t^0;p^1,x^1,t^1)$. We first claim that, for every $r\in\mathbb{R}$, the argument that maximizes $DEU_{r,\delta}$ is decreasing in δ. To see this, consider any pair of parameters (r,δ), and let $a^*\in[0,1]$ be the argument that maximizes $DEU_{r,\delta}$. If $a^*=1$, we are done. Consider then the case of $a^*<1$ and any alternative $a^*<a\le 1$. Given the optimality of a^*, we know that $DEU_{\delta,r}(a)\le DEU_{\delta,r}(a^*)$, i.e., $e^{-\delta t^0}p^0u_r(\omega+(1-a)x^0)+e^{-\delta t^1}p^1u_r(\omega+ax^1)\le e^{-\delta t^0}p^0u_r(\omega+(1-a^*)x^0)+e^{-\delta t^1}p^1u_r(\omega+a^*x^1)$ holds. The latter inequality is equivalent to $p^0u_r(\omega+(1-a)x^0)+e^{-\delta(t^1-t^0)}p^1u_r(\omega+ax^1)\le p^0u_r(\omega+(1-a^*)x^0)+e^{-\delta(t^1-t^0)}p^1u_r(\omega+a^*x^1)$. Now, it is evident that an increase of δ leaves unaffected the first term in both the left and the right hand sides but decreases more significantly the second term of the left hand side, because the function u_r is strictly increasing. Hence, alternative a^* must be preferred to

[1]This is in general different to $A_{\mathcal{T}_3}$. Otherwise, notice that the mapping $\Phi\big(\frac{\mu_\delta+K(A_{\mathcal{T}})\mu_r-K(A_{\mathcal{T}})}{\sqrt{K^2(A_{\mathcal{T}})\sigma_r^2+\sigma_\delta^2+2\rho K(A_{\mathcal{T}})\sigma_r\sigma_\delta}}\big)$ is strictly monotone in ρ, and hence the parameter can be recovered using some other time menu.

alternative a for the larger δ, and the argument maximizing DEU must be a^* or smaller. We have proved our claim. Hence, given $r \in \mathbb{R}$, we can define $K(a, A_{\mathcal{C}}|r)$, $a \in [0,1)$, as the infimum of the values of δ for which any alternative in $[0,a]$ is the DEU maximizer. Hence, the maximizer of $DEU_{r,\delta}$ is below a if and only if δ lies above $K(a, A_{\mathcal{C}}|r)$, i.e., $\Gamma(a, A) = \{(r,\delta) : \delta \leq K(a, A_{\mathcal{C}}|r)\}$, and Claims 1 and 2a follow.

We now study the structure of the thresholds. We start with the case of convex monetary utilities, i.e., $r \leq 0$. Convexity and the fact that $u_r(\omega) = 0$ guarantee that $e^{-\delta t^0} p^0 u_r(\omega + (1-a)x^0) + e^{-\delta t^1} p^1 u_r(\omega + a x^1) \leq e^{-\delta t^0} p^0 [a u_r(\omega) + (1-a) u_r(\omega + x^0)] + e^{-\delta t^1} p^1 [(1-a) u_r(\omega) + a u_r(\omega + x^1)] = e^{-\delta t^0} p^0 (1-a) u_r(\omega + x^0) + e^{-\delta t^1} p^1 a u_r(\omega + x^1) \leq \max\{e^{-\delta t^0} p^0 u_r(\omega + x^0), e^{-\delta t^1} p^1 u_r(\omega + x^1)\}$. Hence, only alternatives 0 or 1 can be the maximizers of $DEU_{r,\delta}$. Thus, for every menu $A_{\mathcal{C}}$ there is a unique threshold $K(a, A_{\mathcal{C}}|r) \in \mathbb{R}$, independent of a, that corresponds to the δ that, given r, equalizes the DEU value of 0 and 1. This value is $\frac{1}{t^1 - t^0} \log \left[\frac{p^1 u_r(\omega + x^1)}{p^0 u_r(\omega + x^0)}\right]$, that can also be written as $\frac{1}{t^1 - t^0} \log \frac{p^1}{p^0} + K(A_{\mathcal{T}}|r)$, with $K(A_{\mathcal{T}}|r)$ referring to the hypothetical time menu in which prizes x^0 and x^1 are offered at periods t^0 and t^1, without considering the probability of these prizes. Proposition 2 argued that $K(A_{\mathcal{T}}|r)$ is strictly decreasing, and hence $K(a, A_{\mathcal{C}}|r)$ is also strictly decreasing whenever $r \leq 0$.

We now analyze strictly concave utilities, $r > 0$. We start by claiming that the threshold $K(a, A_{\mathcal{C}}|r)$ is decreasing for every $a > \bar{e}$, and increasing for every $a < \bar{e}$. We start with the former, assuming by contradiction that $0 < r < r'$ but $K(a, A_{\mathcal{C}}|r) < K(a, A_{\mathcal{C}}|r')$ for some $a > \bar{e}$. Using continuity and the definition of the thresholds, there must exist δ^* with $K(a, A_{\mathcal{C}}|r) < \delta^* < K(a, A_{\mathcal{C}}|r')$ such that the maximizer for DEU_{r,δ^*} is a^*, with $\bar{e} < a^* < a$. Consider any $a' > a^*$. It must be $DEU_{r,\delta^*}(a^*) \geq DEU_{r,\delta^*}(a')$, i.e., $e^{-\delta^* t^0} p^0 u_r(\omega + (1 - a^*)x^0) + e^{-\delta^* t^1} p^1 u_r(\omega + a^* x^1) \geq e^{-\delta^* t^0} p^0 u_r(\omega + (1-a')x^0) + e^{-\delta^* t^1} p^1 u_r(\omega + a' x^1)$. Dividing both terms by the positive constant $p^0 e^{-\delta^* t^0} + p^1 e^{-\delta^* t^1}$ and denoting $p = \frac{p^0 e^{-\delta^* t^0}}{p^0 e^{-\delta^* t^0} + p^1 e^{-\delta^* t^1}}$, the former expression can be written as $p u_r(\omega + (1-a^*)x^0) + (1-p) u_r(\omega + a^* x^1) \geq p u_r(\omega + (1-a')x^0) + (1-p) u_r(\omega + a' x^1)$. Hence, the comparison of these two alternatives is equivalent to that of a risk menu, with a^* corresponding to alternative $0_{\mathcal{R}}$ and a' to alternative $1_{\mathcal{R}}$. Hence, since a^* is preferred at (r, δ^*), we know from Proposition 1 that

a^* will also be preferred at (r', δ^*) because $r' > r$.[2] Thus, the maximizer of DEU_{r',δ^*} cannot be above a^*. This contradicts the definition of $K(a, A_{\mathcal{C}}|r')$. That is, the threshold must be decreasing whenever $a > \bar{e}$. Given that the family $\{u_r\}$ is strictly ordered by concavity, the threshold must be strictly decreasing. The proof that the threshold is strictly increasing whenever $a < \bar{e}$ is analogous and thus omitted.

Consider now $0 < \underline{a} < \bar{e} < \bar{a} < 1$ and denote by $\delta^{\bar{e}}$ the value of δ that makes indifferent all the alternatives when $r = 0$. From the previous reasoning, whenever $r > 0$, $K(\underline{a}, A_{\mathcal{C}}|r)$ (resp., $K(\bar{a}, A_{\mathcal{C}}|r)$) is above (resp., below) $\delta^{\bar{e}}$. Thus, for any given $\delta > \delta^{\bar{e}}$ (resp., $\delta < \delta^{\bar{e}}$), a dominating change in the conditional distribution of r creates an increase in the conditional mass of the set of values of r that lie above the inverse of threshold $K(\underline{a}, A_{\mathcal{C}}|r)$ (resp., $K(\bar{a}, A_{\mathcal{C}}|r)$). Claim 2b then follows.

We now prove Claim 3a. We know that the choice belongs to $[0, \bar{e}]$ if and only if $\delta > K(\bar{e}, A_{\mathcal{C}}|r)$. We can reproduce the analysis of Proposition 2 for the case of expansions in the conditional distribution of δ, and thus, Claim 3a follows. To show Claim 3b, consider any sequence of values $\{a^n\}$, with $a^n > \bar{e}$, such that $\lim_n a^n = \bar{e}$. We know that the choice belongs to $[0, a^n]$ if and only if $\delta > K(a^n, A_{\mathcal{C}}|r)$. Given that for every a^n the threshold is strictly decreasing, we can invert these maps and then reproduce the analysis of Proposition 2 for the case of expansions in the conditional distribution of r, and thus, Claim 3b follows. ■

Proof of Proposition 6: We start by showing that $\theta_r \equiv (\mu_r, \sigma_r)$ is identified. Assume, on the contrary, that this is not the case: there exists θ' and θ^* in Θ with $\theta'_r \neq \theta^*_r$ such that the distribution of the data is the same under both parameters. Let $\log \widetilde{R} \equiv \log \frac{x^1}{x^0} + \log \frac{p^1}{p^0}$ and $k \equiv t^1 - t^0$. Note that, under the assumption that $x^0/\omega \to 0$ and $x^1/\omega \to 0$, the probability of corner allocations, conditional on $r > 0$, converges to zero. Consequently, the probability of observing corner allocations $a \in \{0, 1\}$ is given by $\mathcal{P}_{\theta'}\left(a \in \{0, 1\}, A_{\mathcal{C}}\right) = \mathcal{P}_{\theta'}\left(r \leq 0\right) = \Phi\left(-\mu'_r/\sigma'_r\right)$ which, by the assumption of no-identification, is equal to $\Phi\left(-\mu^*_r/\sigma^*_r\right)$. It follows that $\mu'_r/\sigma'_r = \mu^*_r/\sigma^*_r$. Consequently, both $\mu'_r \neq \mu^*_r$ and $\sigma'_r \neq \sigma^*_r$ must hold. Otherwise, $\theta'_r = \theta^*_r$. Now, using the first order condition of the problem, we

[2] Notice that we are maintaining δ^* constant because this value is part of the definition of the lotteries.

have:

$$\mathbb{E}_{\theta'}\left[\log\left(\frac{ax_m^1+\omega}{(1-a)\,x_m^0+\omega}\right)|r>0\right] = -\mathbb{E}_{\theta'}\left[\frac{\delta}{r}|r>0\right]k_m + \mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right]\log\widetilde{R}_m,$$

for any menu $A_{\mathcal{C},m}$. Combining the corresponding expressions for $A_{\mathcal{C},a}$ and $A_{\mathcal{C},b}$, and using the assumption that $k_a = k_b$, we get:

$$\mathbb{E}_{\theta'}\left[\Delta c|r>0\right] = \mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right]\left(\log\widetilde{R}_a - \log\widetilde{R}_b\right), \tag{A.7}$$

with

$$\Delta c \equiv \log\left(\frac{ax_a^1+\omega}{(1-a)\,x_a^0+\omega}\right) - \log\left(\frac{ax_b^1+\omega}{(1-a)\,x_b^0+\omega}\right).$$

Since the model is not identified, it must also be the case that $\mathbb{E}_{\theta'}\left[\Delta c|r>0\right] = \mathbb{E}_{\theta^*}\left[\Delta c|r>0\right]$. From equation (A.7), it follows that:

$$\left(\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] - \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right]\right)\left(\log\widetilde{R}_a - \log\widetilde{R}_b\right) = 0.$$

Since $\widetilde{R}_a \neq \widetilde{R}_b$ by assumption, the previous expression implies that $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] = \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right]$. Using the fact that r follows a normal distribution, we can write $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right]$ as:

$$\begin{aligned}\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] &= \frac{1}{1-\Phi\left(-\mu_r'/\sigma_r'\right)}\int_0^\infty \frac{1}{x}\left(\frac{1}{\sigma_r'}\right)\phi\left(\frac{x-\mu_r'}{\sigma_r'}\right) \\ &= \frac{1}{\Phi\left(\mu_r'/\sigma_r'\right)\sigma_r'}\int_0^\infty \frac{1}{z}\phi\left(z-\mu_r'/\sigma_r'\right)dz,\end{aligned}$$

where the second line uses the change of variable $z \equiv x/\sigma$ and the symmetry of the Normal CDF . Finally, for θ^* we must also have:

$$\mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right] = \frac{1}{\Phi\left(\mu_r^*/\sigma_r^*\right)\sigma_r^*}\int_0^\infty \frac{1}{z}\phi\left(z-\mu_r^*/\sigma_r^*\right)dz.$$

Since $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] = \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right]$ and $\mu_r'/\sigma_r' = \mu_r^*/\sigma_r^*$, the previous two expressions imply that $\sigma_r' = \sigma_r^*$, arriving to a contradiction. Hence, θ_r is identified.

The next step is to show that $\theta_\delta \equiv (\mu_\delta, \sigma_\delta, \rho)$ are identified. Fix $\theta_r \equiv (\mu_r, \sigma_r)$ and assume, on the contrary, that θ_δ is not identified: there exists θ_δ' and θ_δ^* in Θ such that $\theta_\delta' \neq \theta_\delta^*$ and $\mathcal{P}_{\theta_\delta'}\left(a=0, A_{\mathcal{C}}\right) = \mathcal{P}_{\theta_\delta^*}\left(a=0, A_{\mathcal{C}}\right)$. From Proposition 5, this equality implies: $\mathcal{P}_{\theta_\delta'}\left(a=0, A_{\mathcal{C}}|r\leq 0\right) = \mathcal{P}_{\theta_\delta'}\left(K(0, A_{\mathcal{C}}|r)\leq\delta|r\leq 0\right)\mathcal{P}_{\theta_r}\left(r\leq 0\right)$, which is equal to

$\mathcal{P}_{\theta^*_\delta}\left(K(0,A_{\mathcal{C}}|r)\leq\delta|r\leq 0\right)\mathcal{P}_{\theta_r}\left(r\leq 0\right)$. This equality implies:

$$\int_{-\infty}^{0}\left\{\Phi\left(\frac{K\left(0,A_{\mathcal{C}}|r\right)-\mu'_{\delta|r}}{\sigma'_{\delta|r}}\right)-\Phi\left(\frac{K\left(0,A_{\mathcal{C}}|r\right)-\mu^*_{\delta|r}}{\sigma^*_{\delta|r}}\right)\right\}dr=0,$$

At this stage, identification of θ_δ from convex menus is analogous to its identification using time menus. We can thus use assumption (b) and the same steps used in the second part of the proof of Proposition 3 to arrive to a contradiction. ■

Proof of Proposition 7: We start discussing menus in which $t^1\to t^0$, that can be seen as risk problems only. In these menus, the choice is determined by the optimization of the objective function $p^0\frac{(\omega+(1-a)x^0)^{1-r}-\omega^{1-r}}{1-r}+p^1\frac{(\omega+ax^1)^{1-r}-\omega^{1-r}}{1-r}$. Whenever $r\leq 0$, there are two cases. First, if $p^1x^1\geq p^0x^0$ the choice is $a=1$. Second, if $p^1x^1<p^0x^0$, there is $r^*<0$ such that whenever $r\leq r^*$ the choice is $a=1$ and whenever $r>r^*$ the choice is $a=0$. Now, whenever $r>0$, the solution is interior and choices form a continuous mapping from the corner with larger expectation towards the point $\bar{e}=\frac{x^0}{x^0+x^1}$, which is the limit of choices when $r\to\infty$. Thus, we can consider one such problem, say, one such that $p^1x^1>p^0x^0$. The observed mass of alternative $a=1$, that we denote by q_1, must be equal to $\Phi(\frac{0-\mu_r}{\sigma_r})$. Now, let a be any value in $(\bar{e},1)$ and denote by r_a the value above which the optimal choice falls below a. Denote by q' the observed choice probability below a, that must be equal to $1-\Phi(\frac{r_a-\mu_r}{\sigma_r})$. This allows to obtain parameters μ_r and σ_r: $\sigma_r=\frac{r_a}{1-\Phi^{-1}(q_1)-\Phi^{-1}(q')}$ and $\mu_r=-\sigma_r\Phi^{-1}(q_1)=-\frac{r_a\Phi^{-1}(q_1)}{1-\Phi^{-1}(q_1)-\Phi^{-1}(q')}$

The rest of the parameters can be identified as follows. As commented in the proof of Proposition 4, for any given ω, the use of large payouts is equivalent to use the case $\omega\to 0$ and, in what follows, we assume large payouts. Whenever $p_1\to p_0$, the probability of selecting options below $\frac{1}{2}$ is the probability that δ is above $K(A_{\mathcal{T}})(1-r)$.[3] One can then reproduce the proof of Proposition 4 replacing the mass of $0_{\mathcal{T}}$ for the cumulative mass below $\frac{1}{2}$. ■

[3]Note that unlike in the case of Proposition 4, the probability obtained here is exact.

Appendix B. Hybrid Menus

In a hybrid menu, each of the two alternatives corresponds to a two state-contingent lottery, with the safer lottery awarded earlier in time. Formally, $A_{\mathcal{H}} = \{0_{\mathcal{H}}, 1_{\mathcal{H}}\}$ with $0_{\mathcal{H}} = ([p, 1-p; x_1^0, x_2^0], t^0)$ and $1_{\mathcal{H}} = ([p, 1-p; x_1^1, x_2^1], t^1)$ such that $x_1^1 > x_1^0 > x_2^0 > x_2^1$, $p \in (0,1)$ and $t^0 < t^1$. The analysis of hybrid menus is analogous to that of time menus by conditioning again on parameter r. For any given hybrid menu $A_{\mathcal{T}}$ and any value of r, there exists a menu-dependent constant $K(A_{\mathcal{H}}|r) \in \mathbb{R}$ such that alternative $0_{\mathcal{H}}$ is selected if and only if $\delta \geq K(A_{\mathcal{H}}|r)$.[4] As a result, the choice probability of alternative $0_{\mathcal{H}}$ is:

$$\mathcal{P}_f(0_{\mathcal{H}}, A_{\mathcal{H}}) = 1 - \int_r F_{\delta|r} K(A_{\mathcal{H}}|r) f^r(r)\mathrm{d}r.$$

The effect of shifts and spreads of δ are trivially understood from this structure, by applying the logic of Proposition 2. Understanding the effect of r requires some caution, since the ratio of expected utilities is an object that may be difficult to tame. Fortunately, it can be seen that for standard families of monetary utilities (e.g., CRRA and CARA), the threshold $K(A_{\mathcal{H}}|r)$ is decreasing in r as long as $\delta > 0$.

[4]It is easy to see that $K(A_{\mathcal{H}}|r) = \frac{1}{t^1 - t^2} \log \left[\frac{p u_r(x_1^0) + (1-p) u_r(x_2^0)}{p u_r(x_1^1) + (1-p) u_r(x_2^1)} \right]$.

Appendix C. Numerical Evaluation of Choice Probabilities

DMPL. Given a value for parameter vector $\theta \in \Theta$, computation of the log-likelihood defined in Section 4 requires computing $\mathcal{P}_\theta\left(0_m, A_m\right)$ for each menu A_m in the dataset $\mathcal{O}$. In general, this requires evaluating a double integral numerically. Given the structure of the R-DEU model, this can be done efficiently using Quasi-Monte Carlo (QMC) methods: begin by discretizing the support of f_θ it in N_{QMC} nodes $\{r_k, \delta_k\}_{k=1}^{N_{QMC}}$ using low-discrepancy sequences. Let $\mathcal{I}\left(0_m, A_m | r_k, \delta_k\right)$ denote an indicator function that takes value of one when $DEU_{r_k,\delta_k}\left(0_m\right) > DEU_{r_k,\delta_k}\left(1_m\right)$, and zero otherwise. For large enouh N_{QMC}, we have

$$\mathcal{P}_\theta\left(0_m, A_m\right) \approx \frac{V}{N_{QMC}} \sum_{k=1}^{N_{QMC}} \mathcal{I}\left(0_m, A_m | r_k, \delta_k\right) f_\theta\left(r_k, \delta_k\right),$$

where $V \equiv \int_r \int_\delta dr d\delta = \left(\overline{r} - \underline{r}\right)\left(\overline{\delta} - \underline{\delta}\right)$ is a normalization constant. We can control the accuracy of the approximation by increasing the number of nodes N_{QMC}. Importantly, the indicator function $\mathcal{I}\left(0_m, A_m | r_k, \delta_k\right)$ is independent of θ. It follows that, to compute the maximum-likelihood estimator of θ, this indicator function needs to be computed only once before starting the search of the maximizer, reducing dramatically the estimation time.

We can also use the results in the paper, together with the assumption that f follows a bivariate normal distribution, to improve the estimation algorithm further. For risk menus, Proposition 1 implies $\mathcal{P}_\theta\left(0_m, A_{\mathcal{R},m}\right) = 1 - \Phi\left(\left(K\left(A_{\mathcal{R},m}\right) - \mu_r\right)/\sigma_r\right)$. Computation of $K\left(A_{\mathcal{R},m}\right)$ requires solving a non-linear equation numerically for each risk menu. However, these thresholds are independent of θ so we only need to compute them once before estimation. For time menus, Proposition 2 simplifies the double integral characterizing $\mathcal{P}_\theta\left(0_m, A_{\mathcal{T},m}\right)$ into the following following single-valued integral:

$$\mathcal{P}_\theta\left(0_m, A_{\mathcal{T},m}\right) = 1 - \int_r \Phi\left(\frac{K\left(A_{\mathcal{T},m} | r\right) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right) \phi\left(\frac{r - \mu_r}{\sigma_r}\right) dr,$$

with $\mu_{\delta|r} \equiv \mu_\delta + \rho\frac{\sigma_\delta}{\sigma_r}\left(r - \mu_r\right)$ and $\sigma_{\delta|r} \equiv \sigma_\delta\sqrt{1 - \rho^2}$. This simpler integral can also be evaluated numerically using QMC methods, as discussed before.

Convex Budgets. Computation of the log-likelihood function with convex budgets, as defined in Section 6, is computationally more demanding as it requires evaluating $\mathcal{P}_\theta\left(a \in \alpha^s, A_m\right)$ for each of the M menus in the dataset and each of the S options in which the choice set is discretized. As in the iid-additive RUM, we can proceed by rounding each observed allocation a to the midpoint of the option α^s for which $a \in \alpha^s$. This results in S possible observed allocations in the data: $\overline{\alpha}_1 = 0$, $\overline{\alpha}_2 = \left(a_2 + a_3\right)/2$, ..., $\overline{\alpha}_{S-1} = \left(a_{S-1} + a_S\right)/2$ and $\overline{\alpha}_S = 1$. Let $\mathcal{I}\left(\overline{\alpha}, A_m | r_k, \delta_k\right)$ denote an indicator function that takes value of 1 when $DEU_{r_k,\delta_k}\left(\overline{\alpha}\right) \geq DEU_{r_k,\delta_k}\left(\overline{\alpha}^*\right)$ in menu A_m for all $\overline{\alpha}^* \in \{\overline{\alpha}_1, \ldots, \overline{\alpha}_S\}$, and zero otherwise. We can then use the numerical approximation $\mathcal{P}_\theta\left(a \in \alpha^s, A_m\right) \approx \left(V/N_{QMC}\right) \sum_{k=1}^{N_{QMC}} \mathcal{I}\left(\overline{\alpha}, A_m | r_k, \delta_k\right) f_\theta\left(r_k, \delta_k\right)$. As in the DMPL case, the indicator function $\mathcal{I}\left(\overline{\alpha}, A_m | r_k, \delta_k\right)$ is independent of θ and can be pre-computed before maximization of the log-likelihood function. This method allows for flexible specifications of f and can be easily extended to more general models, as illustrated in Appendix D.

Alternatively, we can exploit the results from Proposition 5 and the assumption that f is normal to compute $\mathcal{P}_\theta\left(\left[a_s, a_{s+1}\right], A_m\right)$ as the single-variable integral:

$$\int_r \left[\Phi\left(\frac{K\left(A_{s,\mathcal{C}}|r\right) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right) - \Phi\left(\frac{K\left(A_{s+1,\mathcal{C}}|r\right) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right)\right] \phi\left(\frac{r - \mu_r}{\sigma_r}\right) dr,$$

where the thresholds $K\left(A_{s,\mathcal{C}}|r\right)$ are defined in Section 5 and are independent of θ.

Appendix D. Extensions

The methods introduced in this paper can be extended to allow for additional behavioral features and alternative distributions for the random parameters. We illustrate this in this section using convex budget data from the experimental design in Andreoni and Sprenger (2012a). This dataset is similar to the one used in Section 5 and features 97 subjects facing 45 convex menus with certain payoffs. The main difference with the data used for the baseline analysis in the main text is that some of the menus feature payoffs in the present, which allows estimation of present bias in discounting.

Table 1 summarizes the results of this exercise. To make comparison across models with different distributional assumptions feasible, we report the median and inter-quantile range (IQR) of the estimated distributions. The second column of Table 1 summarizes the baseline estimates obtained using the R-DEU model following the procedure used in Section 6. The estimated curvature of the utility function is statistically zero, similar to the results obtained by Andreoni and Sprenger (2012b), which could be due to the lack of variation in payoff probabilities in this dataset, indicating that the curvature captures only intertemporal substitution. The estimated median of the annual discount rate is small and statistically close to zero. Nevertheless, the model estimates a large degree of heterogeneity in both parameters.

The third column of Table 1 shows the results using Quasi-Monte Carlo (QMC) methods, as discussed in Section C of this Appendix. The results are virtually identical to those obtained using the baseline algorithm developed in the paper. Nevertheless, it takes four times longer to estimate the model using QMC methods, confirming the benefits of exploiting the economic structure of the problem. Despite this, the QMC method is useful for estimating the model with alternative distributions and behavioral features, as discussed below.

The fourth column of Table 1 shows the results for a "constrained" version of the RDE model where δ follows a normal distribution truncated at zero, ruling out the possibility of preference for the future. In this case, we use a Gaussian copula to allow for correlation between r and δ. The estimated median of the annual discount rate is larger and statistically different from zero. The median curvature of the utility function remains

TABLE 1. Estimated Risk and Time Preferences in Andreoni and Sprenger (2012a)

	R-DEU			Hyperbolic Discounting	
	Baseline	*QMC*	*Constrained*	*Unconstrained*	*Constrained*
Median (r)	−0.051 [0.084]	−0.050 [0.080]	−0.037 [0.039]	−0.039 [0.071]	−0.038 [0.038]
IQR (r)	0.718 [0.166]	0.715 [0.152]	0.347 [0.072]	0.665 [0.122]	0.348 [0.070]
Median (δ)	0.043 [0.157]	0.037 [0.151]	0.504 [0.058]	0.072 [0.148]	0.490 [0.052]
IQR (δ)	2.317 [0.335]	2.310 [0.305]	0.784 [0.087]	2.217 [0.271]	0.764 [0.079]
Median (β)	–	–	–	0.991 [0.004]	0.898 [0.008]
IQR (β)	–	–	–	0.004 [0.005]	0.247 [0.056]
Cor (r, δ)	−0.306 [0.071]	−0.302 [0.068]	0.020 [0.056]	−0.284 [0.067]	0.017 [0.060]
Cor (r, β)	–	–	–	−0.002 [2.605]	−0.078 [0.459]
Cor (δ, β)	–	–	–	−0.001 [0.479]	0.134 [0.356]
$\mathcal{L}$	−1.709	−1.709	−1.835	−1.709	−1.833

NOTES.- The table reports estimated moments of the distributions of risk aversion (r), discounting (δ), and present bias (β) estimated at the population level. The second column reports the results obtained using the R-DEU model with same methodology and assumptions as in Section 6. The third column shows the results using Quasi-Monte Carlo methods (QMC), as discussed in Appendix C. The fourth column shows the results using QMC and assumes that δ follows a normal distribution truncated at zero. The fourth column shows the results for a model extended to allow for present bias under the assumption that the (r, δ, β) follow a multivariate normal distribution. The last column shows the results assuming that δ follows a normal distribution truncated at zero and β a beta distribution. Standard errors for each MLE are shown in brackets and are clustered at the individual level.

close to zero, and the IQR of both parameters is now lower. As expected, restricting the domain of δ reduces the fit to the data, as reflected in the resulting log-likelihood.

The last two columns of Table 1 show the results for the model extended to allow present bias in discounting so that the discount factor in the model is $\beta e^{-\delta}$ when $t^0 = 0$, and $e^{-\delta}$ otherwise. The first of the two columns shows the results for an "unconstrained" model that assumes parameters r, δ and β follow a multivariate normal distribution with an arbitrary correlation matrix. The results in this case are very similar to those obtained in the R-DEU model, indicating a low degree of present bias. The second column shows the results for a "constrained" model where r follows a normal distribution, δ follows a normal distribution truncated at zero, and β follows a beta distribution with support on the unit interval. The estimated distributions of r and δ are similar to those obtained for the constrained R-DEU model, but now the median present bias is statistically different from 1. Nevertheless, the improvement in fit is relatively low compared to the constrained R-DEU model.

Appendix E. Baseline Wealth

We now briefly comment on the role of ω with CRRA monetary utilities. It is immediate to see that in risk menus $A_{\mathcal{R}}$ such that $K(A_{\mathcal{R}}) \neq 0$, $K(A_{\mathcal{R}})$ is strictly increasing (resp., decreasing) in ω whenever $K(A_{\mathcal{R}}) > 0$ (resp., $K(A_{\mathcal{R}}) < 0$).[5] Consequently, ceteris paribus, the alternative with larger expected value will be chosen more often. In time menus $A_{\mathcal{T}}$, every threshold $K(A_{\mathcal{T}}|r)$ converges monotonically to the constant $K(A_{\mathcal{T}}|0)$ as ω increases. The conditional behavior of every r becomes more aligned with the conditional choices of $r = 0$. That is, ceteris paribus, the more risk-averse (resp., lover) individuals will choose more often the present (resp., future) option. Similarly, in convex menus $A_{\mathcal{C}}$, every threshold map $K(a, A_{\mathcal{C}}|r)$ converges monotonically to the constant map $K(a, A_{\mathcal{C}}|0)$ as ω increases. The conditional behavior of every r becomes more aligned with the conditional choices of $r = 0$ (with interior solutions vanishing).

Since in actual practice it is often assumed zero levels of background wealth, it is interesting to discuss theoretically the limit model when the baseline wealth tends to zero. From the previous discussion, we know that this limit case would create the best conditions for parameter r kicking in all menus. Interestingly, for the case of time menus $A_{\mathcal{T}}$, as discussed in the proof of Proposition 3, the conditional threshold map becomes piece-wise linear: (a) when $r < 1$, $K(A_{\mathcal{T}}|r) = \frac{\log \frac{x^1}{x^0}}{t^1 - t^0}(1 - r) \equiv K(A_{\mathcal{T}})(1 - r) > 0$, where $K(A_{\mathcal{T}})$ is a menu-dependent constant, and (b) when $r \geq 1$, $K(A_{\mathcal{T}}|r)$ becomes null. Hence, for parameters (r, δ), with $r < 1$, the earlier option $0_{\mathcal{T}}$ is preferred to the later option for such parameters if and only if $\frac{\delta}{1-r} \geq K(A_{\mathcal{T}})$. That is, the expression $\frac{\delta}{1-r}$ represents a simple correction of δ based on the risk parameter r that captures completely time considerations. In other words, the behavior of $DEU_{r,\delta}$ is equivalent to the behavior of $DEU_{0,\frac{\delta}{1-r}}$, and if the analyst is willing to entertain the idea that risk aversion above 1 is not crucial or that risk aversion and delay aversion are somewhat independent phenomena for standard values, independent distributions of r and $\frac{\delta}{1-r}$ can be considered. Importantly, behavior for $r \geq 1$ becomes extreme when wealth is negligible, as alternative $0_{\mathcal{T}}$ is always preferred.

To see the role of baseline wealth in the empirical applications studied in the paper, Table 2 compares the estimates under the baseline choice of ω with those obtained by

[5] In the degenerate case where the expected values of both lotteries coincide, we obviously have $K(A_{\mathcal{R}}) = 0$ for all levels of ω.

Table 2. Estimated Risk and Time Preferences when $\omega \approx 0$

	DMPL-AHLR		CB-AS	
	Baseline	$\omega \approx 0$	*Baseline*	$\omega \approx 0$
μ_r	0.781 [0.053]	0.681 [0.032]	0.207 [0.062]	0.095 [0.045]
σ_r	0.895 [0.049]	0.768 [0.039]	0.752 [0.079]	0.562 [0.054]
μ_δ	0.125 [0.008]	0.102 [0.007]	0.339 [0.108]	0.383 [0.110]
σ_δ	0.125 [0.010]	0.116 [0.007]	1.805 [0.124]	1.821 [0.125]
ρ	−0.958 [0.016]	−0.999 [0.001]	−0.164 [0.053]	−0.202 [0.052]

NOTES.- The table reports the risk aversion coefficient and annual discount rate at the population level estimated by the R-DEU model under two different assumptions about the value of integrated wealth ω.

setting ω to a positive value close to zero. This exercise confirms the previous theoretical discussion: the estimated average and standard deviation of risk aversion falls with the value of ω, and the correlation between r and δ increases. Nevertheless, these changes are quantitativelly small. Notice also that the estimated marginal distribution of δ remains practically unchanged after across values of ω.

Appendix F. Example of Risk and Time Tasks in AHLR

Table 3. Example of a Risk Task in AHLR

	Lottery A				***Lottery B***			
Menu	p	Prize	p	Prize	p	Prize	p	Prize
1	0.1	2000	0.9	1600	0.1	3850	0.9	100
2	0.1	2000	0.9	1600	0.1	3850	0.9	100
3	0.1	2000	0.9	1600	0.1	3850	0.9	100
4	0.1	2000	0.9	1600	0.1	3850	0.9	100
5	0.1	2000	0.9	1600	0.1	3850	0.9	100
6	0.1	2000	0.9	1600	0.1	3850	0.9	100
7	0.1	2000	0.9	1600	0.1	3850	0.9	100
8	0.1	2000	0.9	1600	0.1	3850	0.9	100
9	0.1	2000	0.9	1600	0.1	3850	0.9	100
10	0.1	2000	0.9	1600	0.1	3850	0.9	100

NOTES.- Example of a risk task in Andersen et al. (2008). All prizes are displayed in DKK.

Table 4. Example of a Time Task in AHLR

Payoff Alternative	***Payment Option A (Pays amount below in 1 month)***	***Payment Option B (Pays amount below in 7 months)***
1	3000	3075
2	3000	3152
3	3000	3229
4	3000	3308
5	3000	3387
6	3000	3467
7	3000	3548
8	3000	3630
9	3000	3713
10	3000	3797

NOTES.- Example of a risk task in Andersen et al. (2008). All prizes are displayed in DKK.

Appendix G. Convex Menus in Andreoni and Sprenger (2012b)

menuID	t^0	t^1	p^0	p^1	q^0	q^1
1	7	35	1	1	0.20	0.20
2	7	35	1	1	0.19	0.20
3	7	35	1	1	0.18	0.20
4	7	35	1	1	0.17	0.20
5	7	35	1	1	0.16	0.20
6	7	35	1	1	0.15	0.20
7	7	35	1	1	0.14	0.20
8	7	63	1	1	0.20	0.20
9	7	63	1	1	0.19	0.20
10	7	63	1	1	0.18	0.20
11	7	63	1	1	0.17	0.20
12	7	63	1	1	0.16	0.20
13	7	63	1	1	0.15	0.20
14	7	63	1	1	0.14	0.20
15	7	35	1	0.8	0.20	0.20
16	7	35	1	0.8	0.19	0.20
17	7	35	1	0.8	0.18	0.20
18	7	35	1	0.8	0.17	0.20
19	7	35	1	0.8	0.16	0.20
20	7	35	1	0.8	0.15	0.20
21	7	35	1	0.8	0.14	0.20
22	7	63	1	0.8	0.20	0.20
23	7	63	1	0.8	0.19	0.20
24	7	63	1	0.8	0.18	0.20
25	7	63	1	0.8	0.17	0.20
26	7	63	1	0.8	0.16	0.20
27	7	63	1	0.8	0.15	0.20
28	7	63	1	0.8	0.14	0.20
29	7	35	0.8	1	0.20	0.20
30	7	35	0.8	1	0.19	0.20
31	7	35	0.8	1	0.18	0.20
32	7	35	0.8	1	0.17	0.20
33	7	35	0.8	1	0.16	0.20
34	7	35	0.8	1	0.15	0.20
35	7	35	0.8	1	0.14	0.20
36	7	63	0.8	1	0.20	0.20
37	7	63	0.8	1	0.19	0.20
38	7	63	0.8	1	0.18	0.20
39	7	63	0.8	1	0.17	0.20
40	7	63	0.8	1	0.16	0.20
41	7	63	0.8	1	0.15	0.20
42	7	63	0.8	1	0.14	0.20
43	7	35	0.5	0.5	0.20	0.20
44	7	35	0.5	0.5	0.19	0.20
45	7	35	0.5	0.5	0.18	0.20
46	7	35	0.5	0.5	0.17	0.20
47	7	35	0.5	0.5	0.16	0.20
48	7	35	0.5	0.5	0.15	0.20
49	7	35	0.5	0.5	0.14	0.20
50	7	63	0.5	0.5	0.20	0.20
51	7	63	0.5	0.5	0.19	0.20
52	7	63	0.5	0.5	0.18	0.20
53	7	63	0.5	0.5	0.17	0.20
54	7	63	0.5	0.5	0.16	0.20
55	7	63	0.5	0.5	0.15	0.20
56	7	63	0.5	0.5	0.14	0.20
57	7	35	0.5	0.4	0.20	0.20
58	7	35	0.5	0.4	0.19	0.20
59	7	35	0.5	0.4	0.18	0.20
60	7	35	0.5	0.4	0.17	0.20
61	7	35	0.5	0.4	0.16	0.20
62	7	35	0.5	0.4	0.15	0.20
63	7	35	0.5	0.4	0.14	0.20
64	7	63	0.5	0.4	0.20	0.20
65	7	63	0.5	0.4	0.19	0.20
66	7	63	0.5	0.4	0.18	0.20
67	7	63	0.5	0.4	0.17	0.20
68	7	63	0.5	0.4	0.16	0.20
69	7	63	0.5	0.4	0.15	0.20
70	7	63	0.5	0.4	0.14	0.20
71	7	35	0.4	0.5	0.20	0.20
72	7	35	0.4	0.5	0.19	0.20
73	7	35	0.4	0.5	0.18	0.20
74	7	35	0.4	0.5	0.17	0.20
75	7	35	0.4	0.5	0.16	0.20
76	7	35	0.4	0.5	0.15	0.20
77	7	35	0.4	0.5	0.14	0.20
78	7	63	0.4	0.5	0.20	0.20
79	7	63	0.4	0.5	0.19	0.20
80	7	63	0.4	0.5	0.18	0.20
81	7	63	0.4	0.5	0.17	0.20
82	7	63	0.4	0.5	0.16	0.20
83	7	63	0.4	0.5	0.15	0.20
84	7	63	0.4	0.5	0.14	0.20

NOTES.- Payoff dates shown in days.