# Supplemental Appendix

## Dynamic Preference "Reversals" and Time Inconsistency

Philipp Strack and Dmitry Taubinsky

# A Mathematical Details for Section 1.1

Denote by $\text{var}_0(\log \mathbb{E}_1[\theta_2]/\theta_1)$ the variance of $\log \mathbb{E}_1[\theta_2]/\theta_1$ conditional on time 0 information.

**Lemma 12.** *Suppose that the agent chooses the allocation $x^0 = 1/2$ at time 0 and the (random) allocation $x^1$ at time 1. Suppose that $\log(x^1/(1-x^1))$ is normally distributed with mean $m$ and variance $v$. Then*

$$\begin{aligned} \mathbb{E}_0[\theta_2] &= \mathbb{E}_0[\theta_1] \\ (\gamma-1)m &= \mathbb{E}_0\left[\log \frac{\mathbb{E}_1[\theta_2]}{\theta_1}\right] + \log(\beta) \\ (\gamma-1)^2 v &= var_0\left(\log \frac{\mathbb{E}_1[\theta_2]}{\theta_1}\right). \end{aligned} \tag{12}$$

**Proof.** Taking first order conditions yields that the optimal effort at time 0 is

$$(\gamma-1)\log\frac{x^0}{1-x^0} = \log\frac{\mathbb{E}_0[\theta_2]}{\mathbb{E}_0[\theta_1]}.$$

The optimal effort at time 1 satisfies

$$(\gamma-1)\log\frac{x^1}{1-x^1} = \log\frac{\mathbb{E}_1[\theta_2]}{\theta_1} + \log(\beta).$$

As $x^0 = 1-x^0$, we have that $\mathbb{E}_0[\theta_2] = \mathbb{E}_0[\theta_1]$. Taking expectations yields that

$$(\gamma-1)m = (\gamma-1)\mathbb{E}_0\left[\log\frac{x^1}{1-x^1}\right] = \mathbb{E}_0\left[\log\frac{\mathbb{E}_1[\theta_2]}{\theta_1}\right] + \log(\beta).$$

Furthermore, we have that

$$(\gamma-1)^2 v = \text{var}_0\left(\log\frac{\mathbb{E}_1[\theta_2]}{\theta_1}\right) \qquad \square.$$

With this Lemma in hand, we now provide calculations for how $\beta$ is identified under the different structural assumptions listed in Table 1.

**Rows 1 and 2: Independent $\theta_2, \theta_1$, revealed at $t = 1$** Suppose that $\theta_1, \theta_2$ are independent and $\log(\theta_1) \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $\log(\theta_2) \sim \mathcal{N}(\mu_2, \sigma_2^2)$. Then (12) becomes

$$\begin{aligned}\exp(\mu_1 + \sigma_1^2/2) = \mathbb{E}_0[\theta_1] = \mathbb{E}_0[\theta_2] &= \exp(\mu_2 + \sigma_2^2/2) \\ (\gamma - 1)m = \mathbb{E}_0[\log(\theta_2/\theta_1)] - \log(\beta) &= \mu_2 - \mu_1 + \log(\beta) \\ (\gamma - 1)^2 v = \sigma_1^2 + \sigma_2^2 &= \sigma_1^2(1 + \sigma_2^2/\sigma_1^2)\,.\end{aligned}$$

The first equation and third equation imply that

$$\mu_2 - \mu_1 = 0.5(\sigma_1^2 - \sigma_2^2) = 0.5\sigma_1^2(1 - \sigma_2^2/\sigma_1^2) = 0.5(\gamma - 1)^2 v \frac{1 - \sigma_2^2/\sigma_1^2}{1 + \sigma_2^2/\sigma_1^2}\,.$$

Plugging into the second equation yields

$$\log(\beta) = (\gamma - 1)m - 0.5(\gamma - 1)^2 v \frac{1 - \sigma_2^2/\sigma_1^2}{1 + \sigma_2^2/\sigma_1^2}\,.$$

In the case of i.i.d. taste shocks, this reduces to

$$\log(\beta) = (\gamma - 1)m\,.$$

The analyst's estimate thus depends on $\sigma_2^2/\sigma_1^2$, which captures what the analyst assumes about how well informed the agent is at time 0 about their time 2 taste-shock relative to their time 1 taste-shock. This ratio captures *both* the variance of the agents' taste shocks as well as the quality of the information about the agents' taste shocks. Setting $\sigma_2^2/\sigma_1^2 = \infty$ captures the case where the analyst assumes that the agent is uninformed about the time 1 preference at time 0. Setting $\sigma_2^2/\sigma_1^2 = 0$ captures the case where the analyst assumes that the agent is uninformed about the time 2 preference at time 0 (and recovers the result we obtained before). If the agent knows equally much about their time 1 and time 2 preferences at time 0, i.e. $\sigma_2^2/\sigma_1^2 = 0$, we get that

$$\log(\beta) = (\gamma - 1)m\,.$$

**Rows 3 and 4: $\theta_1$ learned at $t = 0$ and $\theta_2$ learned in $t = 1$** If $\theta_1$ is learned at time 0 but $\theta_2$ is not, then $\theta_1$ must always take on the value $\theta_1 = \mathbb{E}_1\theta_2$. Then (12) becomes

$$\begin{aligned}\theta_1 &= \mathbb{E}_0[\theta_2] = \exp(\mu_2 + \sigma_2^2/2)\\ (\gamma-1)m &= \mathbb{E}_0[\log(\theta_2/\theta_1)] - \log(\beta) = \mu_2 - \log(\theta_1) + \log(\beta)\\ (\gamma-1)^2 v &= \sigma_2^2\,.\end{aligned}$$

The first equation and third equation imply that

$$\mu_2 - \log(\theta_1) = -0.5\sigma_2^2 = -0.5(\gamma-1)^2 v\,.$$

Plugging this into the second equation yields that

$$\log(\beta) = 0.5(\gamma-1)^2 v + (\gamma-1)m$$

**Rows 5 and 6: $\theta_2$ learned in $t = 0$ and $\theta_1$ learned in $t = 1$** If $\theta_2$ is learned at time 0 but $\theta_1$ is not, then $\theta_2$ must always take on the value $\theta_2 = \mathbb{E}_1\theta_1$. Then (12) becomes

$$\begin{aligned}\theta_2 &= \mathbb{E}_0[\theta_1] = \exp(\mu_1 + \sigma_1^2/2)\\ (\gamma-1)m &= \mathbb{E}_0[\log(\theta_2/\theta_1)] - \log(\beta) = \log(\theta_2) - \mu_1 + \log(\beta)\\ (\gamma-1)^2 v &= \sigma_1^2\,.\end{aligned}$$

The first equation and third equation imply that

$$\mu_1 - \log(\theta_2) = -0.5\sigma_1^2 = -0.5(\gamma-1)^2 v\,.$$

Plugging this into the second equation yields that

$$\log(\beta) = -0.5(\gamma-1)^2 v + (\gamma-1)m$$

**Rows 7 and 8: $\theta_1$ learned in $t = 1$ and $\theta_2$ learned after $t = 1$** Assume that $\log(\theta_1)$ is Normally distributed with mean $\mu$ and variance $\sigma^2$, and assume, without loss of generality, that $\mathbb{E}[\theta_1] = 1$. By assumption we have that $\mathbb{E}_0[\theta_2] = \mathbb{E}_1[\theta_2]$ and hence (12) becomes

$$\begin{aligned}1 &= \mathbb{E}_0[\theta_1] = \exp(\mu_1 + \sigma^1/2)\\ (\gamma-1)m &= \log(\mathbb{E}_0[\theta_2]) - \mathbb{E}_0\left[\log(\theta_1)\right] + \log(\beta)\\ (\gamma-1)^2 v &= \mathrm{var}_0(\log(\theta_1)) = \sigma_1^2\,.\end{aligned}$$

The first and third equation together imply that $\mu = -\sigma^2/2 = -0.5(\gamma-1)^2 v$ and hence

$$\log(\beta) = (\gamma-1)m - [\mu + \sigma^2/2] + \mu = (\gamma-1)m - \sigma^2/2 = (\gamma-1)m - 0.5(\gamma-1)^2 v\,.$$

**Rows 9 and 10: Multiplicative random walk with $\theta_1, \theta_2$ both learned in $t = 1$** Formally, $\theta_2 = \theta_1 \cdot \epsilon_1$, where $\log(\epsilon_1) \sim N(\mu, \sigma)$ is log-Normally distributed and $\theta_1, \theta_2$ are learned by the agent only at the beginning of time 1. Then (12) becomes

$$\begin{aligned} 1 &= \mathbb{E}_0[\epsilon_1] = \exp(\mu + \sigma^2/2) \\ (\gamma-1)m &= \mathbb{E}_0[\log(\epsilon_1)] - \log(\beta) = \mu - \log(\beta) \\ (\gamma-1)^2 v &= \sigma^2\,. \end{aligned}$$

The first and third equation together imply that $\mu = -\sigma^2/2 = -0.5(\gamma-1)^2 v$ and hence

$$\log(\beta) = (\gamma-1)m + 0.5(\gamma-1)^2 v\,.$$

**Rows 11-14: Mulitplicative AR(1), with $\theta_1$ learned in $t = 1$ and $\theta_2$ learned after $t = 1$** Suppose that $\log(\theta_2) = \alpha \log(\theta_1) + \log(\varepsilon)$, where $\log(\theta_1) \sim N(\mu_1, \sigma_1^2)$ and $\log(\varepsilon) \sim N(\mu_\varepsilon, \sigma_\varepsilon^2)$. That is, $\log(\theta_1)$ and $\log(\theta_2)$ form an AR(1) process. The agent learns $\theta_1$ at time 1 and $\varepsilon$ at time 2. The scalar $\alpha$ can be regarded as a parametrization of how much is learned at about time-1 versus time-2 shocks at time 1. For example, $\alpha = 0$ means that nothing is learned about time-2 shocks at time 1, while $\alpha \to \infty$ captures the case where at time 1 the agent mostly learns about time-2 shocks.

Under this process, we have that $\log\theta_2|\theta_1 \sim N(\alpha \log\theta_1 + \mu_\varepsilon, \sigma_\varepsilon^2)$ and $\mathbb{E}_1[\log\theta_2|\theta_1] = \alpha\theta_1 + \mu_\varepsilon + \sigma_\varepsilon^2/2$. Thus, (12) becomes

$$\begin{aligned} \exp(\mu_1 + \sigma_1^2/2) = \mathbb{E}_0[\theta_1] = \mathbb{E}_0[\theta_2] &= \exp(\alpha\mu_1 + \alpha^2\sigma_1^2/2 + \mu_\varepsilon + \sigma_\varepsilon^2/2) \\ (\gamma-1)m &= \mathbb{E}_0[\alpha\log\theta_1 + \mu_\varepsilon + \sigma_\varepsilon^2/2 - \log\theta_1] + \log\beta = (\alpha-1)\mu_1 + \mu_\varepsilon + \sigma_\varepsilon^2/2 + \log\beta \\ (\gamma-1)^2 v &= (\alpha-1)^2\sigma_1^2\,. \end{aligned}$$

The first equality implies that

$$(\alpha^2-1)\mu_1 + \mu_\varepsilon + \sigma_\varepsilon^2/2 = (1-\alpha^2)\sigma_1^2/2.$$

The third equality implies that

$$\sigma_1^2 = (\gamma-1)^2 v/(\alpha-1)^2$$

and thus that

$$(1-\alpha^2)\sigma_1^2/2 = 0.5(\gamma-1)^2 v\frac{1-\alpha^2}{(\alpha-1)^2}$$

Plugging this into the expression for $(\gamma-1)m$ yields

$$\begin{aligned}\log\beta &= (\gamma-1)m - 0.5(\gamma-1)^2 v\frac{1-\alpha^2}{(\alpha-1)^2}\\ &= (\gamma-1)m + 0.5(\gamma-1)^2 v\frac{1+\alpha}{\alpha-1}\end{aligned}$$

Now when $\alpha > 1$, $\beta$ becomes arbitrarily large as $\alpha$ converges to 1 from the right. When $\alpha < 1$, $\beta$ becomes arbitrarily small as $\alpha$ converges to 1 from the left.

# B Relation to Other Technical Results

**Social Choice.** The ordinal efficiency welfare theorem (McLennan, 2002, Carroll, 2010) states that for any lottery that is Pareto efficient given a vector of ordinal preferences, there exist utility functions consistent with the ordinal preferences such that this lottery maximizes the sum of utilities. This result is mathematically equivalent to the special case of Proposition 3 where the analyst only observes the most preferred time-0 alternative.[34] The sharper and more interesting characterizations that we provide for single-peaked and concave preferences in Theorems 1 and 2 do not, to our knowledge, relate to any known results in the social choice literature—although they of course have implications for that literature. For example, they imply that for complete single-peaked preferences, it is not necessary to consider lotteries: an alternative is a maximand of some social welfare function as long as it is not Pareto-dominated by any other alternative. Example 4 shows that this stronger conclusion fails for social choice problems without the single-peaked property.[35]

**Dynamically Consistent Preferences over Acts.** A literature in decision theory has studied the question of when preferences over acts are consistent with EU (e.g. Chapter 8.2 in Strzalecki, 2021). In this literature, the analyst observes any decision-relevant state as well as preferences over acts. This contrasts with our setting where states are unobserved and only preferences over actions—i.e., constant acts—are observed by the analyst. For example, in the context of food choices, the assumption made in this literature would correspond to

[34] Specifically, this is the case for the more general version stated by Carroll (2010). The original version stated by McLennan (2002) imposes a more special structure.

[35] It is perhaps also worth clarifying that to our knowledge and understanding, our results do not have a mathematical connection to the literature on aggregation of time preferences (e.g., Jackson and Yariv, 2015, Millner, 2020).

the analyst observing how hungry the agent is, what type of meal he had last, and whether or not it is a warm day, as well as preferences over strategies that specify at time 0 what the agent will eat in *each* of these observable states. Notably, such a data set—where states and preferences over strategies are observable—is much richer than the data sets collected in the preference reversal literature, which are our objects of study. This decision literature refers to the analog of our no sure direct preference reversals condition on acts as "dynamic consistency" (Axiom 8.6 in Strzalecki, 2021). Imposed over acts this condition is much more restrictive and (together with consequentialism) implies that there is a subjective EU representation of the preference (Theorem 8.10 and Theorem 8.24 in Strzalecki, 2021 and Ghirardato, 2002). This is in contrast to our setting where we show that "no sure direct preference reversal" is, without the restriction to single-dimensional choice sets and single-peaked preferences, not sufficient to ensure the existence of an EU representation.

**Random Utility Models.** In the literature on random utility models, the analyst observes the distribution of *optimal choices* from *all* choice sets at a single point in time (comparable to our time-1 data $(\preceq^1, f)$). The question is what can be learned about the agent's mean utility for the different alternatives. By contrast, we assume that the analyst observes the distribution of *preferences* over a choice set. This data can not be reconstructed from the optimal choices (Fishburn, 1998). The data sets we study, which are based on the types of experimental data collected in practice, are therefore richer. Allowing the analyst to observe the distribution over a complete ranking of all alternatives is equivalent to allowing the analyst to observe a joint distribution of preferred alternatives from all subsets in the random utility literature.[36] While our time-1 data is always consistent with EU, one needs additional conditions to ensure consistency with EU when only the marginal distribution of choices from subsets, but not the joined distribution is observed. A focus of the random utility literature has been to identify such conditions (Block et al., 1959, McFadden and Richter, 1990, Clark, 1996, Gul and Pesendorfer, 2006).

A second difference is that the random utility literature typically makes the "positivity" assumption that each alternative is the most preferred one with positive probability. This is a strong assumption when combined with the assumption of single-peaked preferences, which are the main focus of our paper. Positivity and single-peakedness together imply that the agent ranks the alternatives both in increasing and decreasing order with positive probability. Furthermore, a corollary of our Proposition 1 implies that this assumption is highly consequential, as it implies that the average utility cannot be identified without

[36] Formally, when observing the distribution $f$ over strict rankings, one can infer the probability of choosing $x$ from the set $M \subseteq X$ as $\sum_\omega f_\omega \mathbf{1}_{x \succeq^1_\omega y \forall y \neq x}$.

imposing additional structure on the preference shocks. There is also a thematic, but not mathematical, connection to identifying time preferences in dynamic discrete choice models. See, e.g., Magnac and Thesmar (2002), Abbring and Daljord (2020), Levy and Schiraldi (2020), Mahajan et al. (forthcoming). This generalizes the insight from Alós-Ferrer et al. (2021) who highlight a related identification issue in a setting where the analyst has less information and only observes the marginal distribution of preferences over binary choice sets. They propose to resolve it by inferring cardinal information from response times, which is similar to the additional choice dimension we propose in Section 5.2.The literature on dynamic random utility (e.g. Fudenberg and Strzalecki, 2015, Frick et al., 2019) studies questions that are further removed from ours. We are interested in settings where the agent makes the same choice repeatedly, while that literature studies when a sequence of dynamic choices can be rationalized if the agent's utility function and choice set can change over time. An exception is the case of Bayesian evolving beliefs discussed in Section 6.2 of Frick et al. (2019). Their Proposition 6 concerns a special case of their model which is similar to a special case of our Proposition 3 where preferences over some set of lotteries are observable.

Similarly, our model is different from those analyzed in the literature on preferences for flexibility due to taste uncertainty, as in Ahn and Sarver (2013), where the agent chooses a menu at time-0 and then chooses from that menu at time-1.