Did Harold Zuercher Have Time Separable Preferences?

Jay Lu Yao Luo Kota Saito Yi Xin

(UCLA, Toronto, and Caltech)

August 7, 2024

DSE Conference in Policy Evaluation and Heterogeneity Measurement

University of Wisconsin-Madison

Motivation

- The traditional time additive utility function:

$$\sum_{t=0}^{\infty} \beta^t \mathbb{E}[u(c_t)]$$

  - widely adopted due to mathematical and computational convenience.
- Impose several restrictions:
  - both willingness to substitute across time and willingness to substitute across states are captured by u.
  - indifference to the timing of the resolution of uncertainty.

Motivation

- Models with non-separable time preferences such as Epstein and Zin (1989) allow for a clean separation of time and risk preferences.
- This generality makes their use popular in many fields such as macroeconomics and finance.
  - explain the equity premium puzzle (e.g., Mehra and Prescott (1985) and Bansal and Yaron (2004)).
- The importance of non-separable preferences has been known in the dynamic discrete choice literature.

Motivation

Rust (1994) points out:

- *"A number of experiments have indicated that human decision-making under uncertainty may not always be consistent with the von Neumann-Morgenstern axioms.*
- *In addition, expected-utility models imply that agents are indifferent about the timing of the resolution of uncertain events, whereas human decision-makers seem to have definite preferences over the time at which uncertainty is resolved."*

Our Paper

- We develop theoretical results and empirical tools for analyzing DDC models with non-separable time preferences.
- Our model introduces several novel features:
  - flexibly captures risk and time preferences.
  - allows non-trivial preferences over when uncertainty is resolved.
  - nests the standard time-separable expected utility model.
- An application to the bus engine replacement data in Rust (1987): using separable time preferences leads to biased estimates in a systematic way.

Non-separable Time Preferences

- Recall standard models with separable time preferences:

$$\mathbb{E}_c[(1-\beta)u(c) + \beta \mathbb{E}_{v|c}[v]]$$

Risk is <u>additively separable</u> across time.

- Consider non-separable time preferences:

$$\mathbb{E}_c[\phi\big((1-\beta)u(c) + \beta \phi^{-1}(\mathbb{E}_{v|c}[v])\big)]$$

ϕ is a strictly increasing aggregator function, its curvature captures the agent's attitudes towards how risk is resolved across different time periods.

Parametric Special Cases

- Linear ϕ: reduce to standard separable models.
- CRRA Epstein-Zin:

$$u(c) = c^{1-\rho}, \quad \phi(z) = z^{\frac{1-\alpha}{1-\rho}}.$$

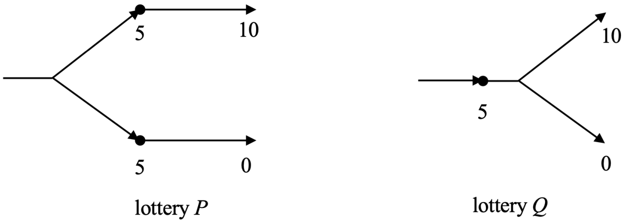
- CARA Epstein-Zin:

$$u\left(c\right) = \rho^{-1}\left(1 - e^{-\rho c}\right), \quad \phi\left(z\right) = \alpha^{-1}\left(1 - \left(1 - \rho z\right)^{\frac{\alpha}{\rho}}\right).$$

- α and ρ characterize risk and intertemporal preferences. When $\alpha = \rho$ in either case, ϕ becomes linear.

Early vs. Late Resolution of Uncertainty

- ϕ allows for greater flexibility in modeling agent's non-trivial preferences over when risk is resolved.



- The two lotteries are the same in the probabilities and outcomes: they differ only in the timing (the uncertainty about period 2 outcome is resolved earlier in P).
- CRRA and CARA: the agent prefers early resolution of uncertainty (i.e., prefers P to Q) iff $\rho \leq \alpha$.

Empirical Model

- Per-period consumption: $c(d_t, x_t, \varepsilon_t) = \pi(d_t, x_t) + \varepsilon_t(d_t)$.
  - d_t: choice
  - x_t: observable state
  - $\varepsilon_t = (\varepsilon_t(1), \varepsilon_t(2), \cdots, \varepsilon_t(J)) \sim G$: indiosyncratic shocks.
- Bellman equation:

$$V(x_t, \varepsilon_t) = \max_d \phi\Big((1-\beta)\, u\big(c\,(d, x_t, \varepsilon_t)\,\big) + \beta\phi^{-1}\big(\mathbb{E}[V(x_{t+1})]\big)\Big).$$

- Ex-ante value: $V(x_{t+1}) = \int V(x_{t+1}, \varepsilon_{t+1}) dG(\varepsilon_{t+1})$.

Value Function Iteration Operator

$$T(V)(x_t) = \int \max_{d} \left\{ \phi\bigg((1-\beta)\, u\big(c\,(d, x_t, \varepsilon_t)\,\big) + \beta\phi^{-1}\big(\mathbb{E}[V(x_{t+1})]\big) \bigg) \right\} dG(\varepsilon_t)$$

- The ex-ante value V is obtained as a solution of the functional equation $T(V) = V$.
- Existence? Uniqueness?

Existence of the Fixed Point

Theorem 1

Suppose v^* *and* v_* *are finite and let* $\mathcal{V}$ *be the set of functions bounded by some* $\underline{v} \leq v_*$ *and* $\bar{v} \geq v^*$*. Then* $T : \mathcal{V} \to \mathcal{V}$ *has a fixed point. Moreover,* $\lim_n T^n(\underline{v})$ *and* $\lim_n T^n(\bar{v})$ *are its smallest and largest fixed points respectively.*

- The proof makes use of lattice theory and Tarski's fixed point theorem.
- Provide a way to compute the smallest and largest value functions by iterating T.
- Week conditions needed for specific parameterizations:
  - CRRA Epstein-Zin: $\mathbb{E}[\varepsilon]$ is finite.
  - CARA Epstein-Zin: $\mathbb{E}\left[e^{-t\varepsilon}\right]$ is finite for all $t \in \mathbb{R}$.

  These conditions are easily satisfied for normal or type I extreme value distributions.

Contraction Mapping Theorem

Theorem 2

T is a contraction mapping if

$$M := \mathbb{E}\left[\max_{d \in D} \sup_{\pi \in [\underline{\pi}, \bar{\pi}], z \in [\underline{v}, \bar{v}]} \psi'_{u(\pi+\varepsilon_d)}(z)\right] < 1$$

where $\psi_y(z) := \phi((1-\beta)y + \beta\phi^{-1}(z))$.

- Simplified formula for specific parameterizations:
  - Linear ϕ (separable time preferences): $\beta < 1$
  - CRRA Epstein-Zin: $\rho \leq \alpha$ and $\beta^{\frac{1-\alpha}{1-\rho}} < 1$
  - CARA Epstein-Zin: $\rho \geq \alpha$ and $\beta^{\frac{\alpha}{\rho}} < 1$

Estimation

- Nested Fixed Point algorithm.
  - Let $\theta = (\alpha, \rho, \theta_\pi)$ be the vector of structural primitives.

$$LL(\theta) = \sum_{i}^{N} \log\left(\prod_{t=1}^{T} \underbrace{Pr(d_{it}|x_{it};\theta)}_{\text{CCP}} \underbrace{Pr(x_{it}|x_{it-1}, d_{it-1})}_{\text{state transition}} \right)$$

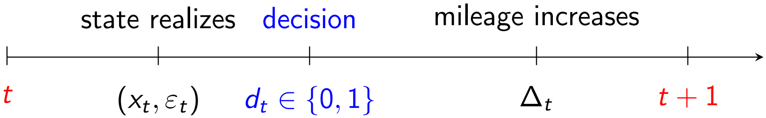
  - Outer loop: search for θ to maximize the LL.
  - Inner loop: given θ, solve for V via iteration.
  - No closed-form solutions for CCP and V (simulation methods needed).

Engine Replacement Model Revisited

- We apply our model to the bus engine replacement decisions originally studied in Rust (1987).
  - 104 buses managed by Harold Zuercher.
  - 10 years of monthly data on bus mileage and engine replacements.

- We extend the model to allow
  - the agent to have non-separable time preferences
  - the agent to be risk averse
  - the agent to earn revenue from operating the bus

Timeline



- x_t: accumulated mileage observed at the beginning of t.
- ε_t: unobserved payoff shock.
- Δ_t: incremental mileage realized within period t.

$$\pi(\Delta_t, d_t, x_t) = \begin{cases} \underbrace{\theta_d \Delta_t}_{\text{revenue}} - \underbrace{RC}_{\text{replacement cost}} & \text{if } d_t = 1 \\ \theta_d \Delta_t - \underbrace{\theta_x x_t}_{\text{maintenance cost}} & \text{if } d_t = 0 \end{cases}$$

Value Functions

- Apply CARA parameterization (to allow negative payoff).
- Choice-specific value functions

$$v(d_t, x_t, \varepsilon_t) = \frac{1}{\alpha}\Bigg[1 - E_{\Delta_t|x_t,d_t}\Bigg\{(1-\beta)\exp\Bigg(-\rho\underbrace{(\pi(\Delta_t, d_t, x_t) + \sigma\varepsilon_t(d_t))}_{\text{current period payoff}}\Bigg)$$

$$+\beta\Bigg(1 - \alpha\underbrace{V(x_{t+1})}_{\text{ex-ante value}}\Bigg)^{\frac{\rho}{\alpha}}\Bigg\}^{\frac{\alpha}{\rho}}\Bigg],$$

where $x_{t+1} = (1 - d_t)x_t + \Delta_t$.

Different Model Specifications

- We estimate four model specifications
  (1) non-separable time preferences (no restrictions on α and ρ);
  (2) separable time preferences with risk aversion ($\alpha = \rho$);
  (3) Rust model with revenue (separable and risk neutral);
  (4) original Rust model (fix $\theta_d = 0$).
- Fix $RC = 8$ (this is the number reported by HZ, see Rust (1987) Table 3), $\beta = 0.9$, and $\varepsilon \sim N(0, 1)$. Estimate $\theta = (\alpha, \rho, \theta_d, \theta_x, \sigma)$.

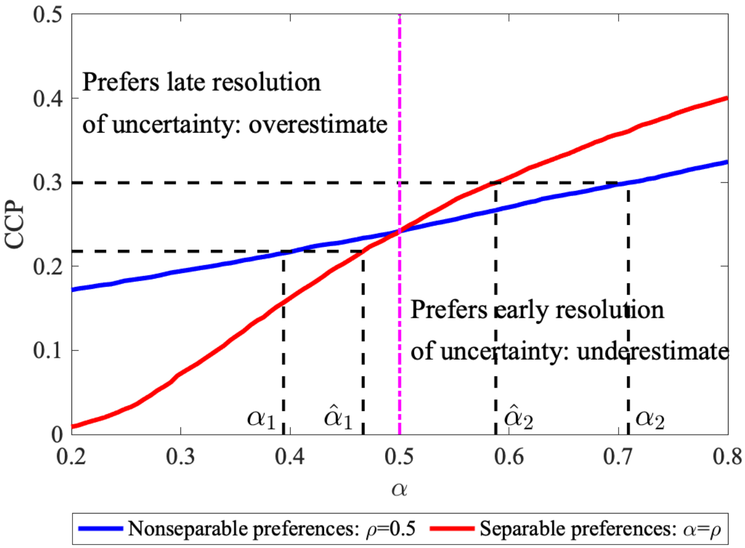
Results

	(1) Nonseparable	(2) Separable	
θ_d	0.0526	0.1019	diff estimates of θ_d
	(0.0073)	(0.0829)	
θ_x	0.1077	0.0329	diff estimates of θ_x
	(0.0122)	(0.0022)	
σ	1.6070	1.5436	
	(0.0491)	(0.0623)	
α	0.1023	0.1457	over estimate α
	(0.5087)	(0.0095)	
ρ	0.5555		prefers late resolution
	(0.0200)		of uncertainty
LL	-299.4404	-300.8139	Reject separable pref

Results

	(1) Nonseparable	(2) Separable	(3) Rust - rev	(4) Rust-orig
θ_d	0.0526	0.1019	0.0001	
	(0.0073)	(0.0829)	(0.0351)	
θ_x	0.1077	0.0329	0.0208	0.0208
	(0.0122)	(0.0022)	(0.0009)	(0.0014)
σ	1.6070	1.5436	1.4883	1.4770
	(0.0491)	(0.0623)	(0.0542)	(0.0566)
α	0.1023	0.1457		
	(0.5087)	(0.0095)		
ρ	0.5555			
	(0.0200)			
LL	-299.4404	-300.8139	-301.4402	-301.6273

Biased Estimates



Counterfactual

- The agent has uncertainty about the incremental mileage:
  - affects current-period revenue
  - affects future maintenance costs
- A subsidy program:
  - helps the agent smooth revenue across periods.
  - the agent receives C each period even if the incremental mileage is 0.
  - similar to the "capacity payment".

Counterfactual

- Agents with separable preferences: $C = \$120$ (almost doubles!)
- Agents with nonseparable preferences: $C = \$61.6$
- Intuition: when the agent prefers late resolution of uncertainty, the original setting (uncertainty in revenue) is less unfavorable.
- Misspecifying agents' preferences as time-separable when it is in fact not leads to misleading policy recommendations.

Main Takeaways

- It's well known that non-separable time preferences is more realistic but challenging to incorporate in DDC models.
- What we've shown:
  - nice theoretical properties: existence, contraction.
  - estimation is fairly straightforward: simulated nested fixed point.
  - applicable to many empirical settings.
  - ignoring nonseparable time preferences may severely bias the estimates and distort policy implications.

Related Literature

- Theory works:
  - Kreps and Porteus (1978), Epstein and Zin (1989, 1991).
  - Lu and Saito (2020): a model of dynamic stochastic choice that incorporates Epstein and Zin.
- Dynamic discrete choice:
  - Rust (1987) and many others.
  - No works on dynamic discrete choice that incorporates non-separable preferences.

Discussion on Identification

- What data variations are needed to empirically distinguish nonseparable preferences from separable preferences?

- A two-states problem: $s_t \in \{0, 1\}$ and $\Delta_t \in \{0, 1\}$.

$$Pr(\Delta_t = 1 | d_t = 1, s_t) = 1,$$

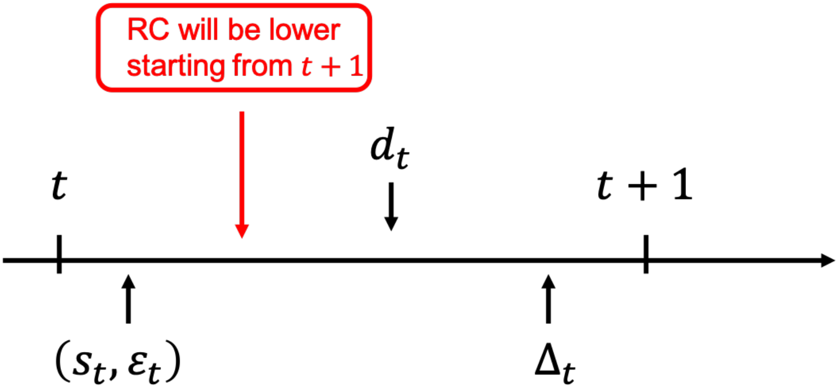
$$Pr(\Delta_t = 1 | d_t = 0, s_t = 0) = 1,$$

$$Pr(\Delta_t = 1 | d_t = 0, s_t = 1) = p, \quad p \in (0, 1)$$

- Future state:

$$s_{t+1} = \max\{(1 - d)s_t + \Delta_t, 1\}.$$

Exogenous Shifters



- This information has <u>no direct effect</u> on the agent's current payoff, but will change the <u>future value</u> $V(1)$.
- Separable preferences: the changes in V for both options cancel out, CCP remains the same.

A Graphical Illustration