Quantifying Uncertainty over the Lifecycle

Nadav Kunievsky

October 31, 2023

Motivation

- The goal of this project is to measure the lifetime uncertainty costs across different social groups in India
- Uncertainty over the future can generate large differences in welfare and can aggregate to large differences in other aggregate variables such as wealth and income
- This is not the first project that aims at capturing uncertainty
  - Cunha and Heckman (circa 2005)
  - Arellano et al. (2022)
  - Blundell et al. (2008)
  - Lucas (1987), Barlevy (2004), Alvarez (2004)

Overview

- In this project, I measure individuals' willingness to pay for resolving all future uncertainty/The value of ex-ante information
- The cost of uncertainty is a function of two components
  - The information structure
  - The utility/loss function and the associated actions
- As these are the two main components that affect uncertainty costs, I want to make as few as possible assumptions on their structure

Overview

- Model
- Uncertainty measure
- Data
- Issues with the data
- Estimation using generative models (Athey et al. (2020); Kaji et al. (2022))

Model

- Let $\Omega = \{\{y_t\}_0^T, \{R_t\}_0^T, \{V_t\}_0^T\}$ be the state space and let $\eta = \{\eta\}_0^T$ be a sequence of signals drawn each period from $P(\eta|\omega)$
- We assume that at period $t = -1$ agents have common prior over sequences $\pi(\Omega)$
- Agents have rational expectations but may have more information than what is available to the researcher
- Agents have the same utility function (may vary with observables).

Model

- Each period τ the DM solves

$$\max_{C_t} E\left[\sum_{t=\tau}^{T} \beta^t u(C_t, V_t) \middle| s_\tau\right]$$

subject to the budget constraint

$$C_t + A_t \leq A_{t-1}R_t + y_t$$

where C_t is consumption, V_t are taste shifters, A_t are assets and R_t are the returns on assets

Parameter of interest

- Let $s^t = (\{y_t\}_t^T, \{R_t\}_t^T, \{V_t\}_t^T, \{\eta_t\}_t^T)$ and $s_t = (\{y_t\}_0^t, \{R_t\}_0^t, \{V_t\}_0^t, \{\eta_t\}_0^t, A_{t-1})$
- Let

$$W_{Uncertainty}(s_\tau) = \max_{\{C_t(s_t), A_t(s_t)\}_t^T, \in \mathcal{C}(s_t) \forall s_{t|} | s_\tau} \mathrm{E}\left[\sum_{t=\tau}^{T} \beta^t u(C_t(s_t), V_t(s_t)) | s_\tau \right]$$

$$\mathcal{C}(s_t) = \left\{ C_t(s_t), A_t(s_t) : C_t + A_t \leq A_{t-1} R_{t-1}(s_t) + y_t(s_t) \right\}$$

and

$$W_{CI}(s_\tau) = \mathrm{E}\left[\max_{\{C_t(s^t)\}_t^T, \{A_t(s^t)\}_t^T \in \mathcal{C}(s^t)} \sum_{t=\tau}^{T} \beta^t u(C_t, V_t) | s_\tau \right]$$

$$\mathcal{C}(s^t) = \left\{ \{C_t\}_t^T, \{A_t\}_t^T : \forall t \in \{t, ..., T\}, C_t + A_t \leq A_{t-1} R_{t-1}(s^t) + y_t(s^t) \right\}$$

Parameter of interest

- The normalized cost of uncertainty, given realization s_t, is

$$H(s_T) = \frac{W_{CI}(s_T) - W_{Uncertainty}(s_T)}{u_c(C_T, V_T)}$$

- This measure captures in dollars how much agents are willing to pay in order to resolve all uncertainty
- Usually it's the case we can't observe s_t, but we observe part of the information set X_t
- Therefore, we can identify and estimate $E[H(s_t)|X_t]$, due to the law of iterated expectations, the time separability of the utility function and rational expectations assumption

Alternative Measures - Group level

- **Current Compensation Compensation** - Find C such that

$$H_{compensation}(s_T) = W_{CI}(s_T) - (W_{Uncertainty}(s_T) - u(C_T, V_T)) + u(C, V_T) = 0$$

- Captures the amount of current period consumption we need to make the agent indifferent
- Issues:
  - Sensitive to the current utility function
  - Not clear what to do if we allow for different consumption types
- We then calculate $E[C|x]$ for each X of interest

Alternative Measures - Group Level

- Let $W_{CI}^{K}(s_\tau)$ be the value of complete information and asset level K
- **Cost Measure -** δ such that

$$H_{Cost}(s_\tau) = W_{CI}^{A_0(S_\tau)-\delta}(s_\tau) - W_{Uncertainty}(s_\tau) = 0$$

- Captures the amount of current period assets we can take from the agent to make her indifferene
- Issues:
  - In practice we solve for δ that satisfies

$$E[W_{CI}^{A_0-\delta}(s_\tau) - W_{Uncertainty}(s_\tau)|A_0, X] = 0$$

    which gives us the average cost for the observed group

  - Hard to compute

We then calculate $E[\delta|x]$ for each X of interest

Data - CPHS

- We use the Center for Monitoring Indian Economy's Consumer Pyramids Household Survey (CPHS)
- This is the largest Household (HH) survey in the world, covering around 200K HH across India
- The survey assures that every HH is interviewed every four-month on their income sources, expenditure, work and employment, material status, and demographics.
- The period covered is 2014-2022; in my setup, I restrict attention to 2015-2019 (60 months)
- Today, I am using a smaller sample of HH that answered consistently for the 60 months

Identification

- Identification of the target parameter requires
  - Identification of the flow utility functions
    - Follows from nonparametric identification of the Euler equation (Escanciano et al. (2021))
  - Identification of the joint distribution $\pi(\{Y_t\}, \{R_t\}, \{C_t\}, \{V_t\}, A_0)$
    - In theory, straightforward, in practice data limitations requires making additional assumptions

Identification of the Marginal Utility - Intuition

- Assume we observe the joint distribution of $\pi\{\{C_t\}_0^T, \{R_t\}_0^T, \{V_t\}_0^T\}$
- Identification of the flow utility function builds on the Euler equation and the results in Escanciano et al. (2020) (up to multiplicative and additive constants)

Identification of the Marginal Utility - Intuition

- Let $q = (C, V)$ and $q' = (C', V')$
- As we assumed that the preferences are time separable, the Euler equation implies
$$u_c(q) = \beta \mathrm{E}[u_c(q')R' \mid q]$$
- We can rewrite the equation as
$$u_c(q) = \beta \int_q u_c(q')\psi(q, q')dq'$$
where $\psi(q, q') = E[R' \mid q, q']p(q' \mid q)$

Identification of the Marginal Utility - Intuition

- To see the intuition behind the identification result of Escanciano et al. (2020), consider the finite case in which $q \in \{q_1, ..q_K\}$. Then the Euler equation is written as

$$u_c(q_i) - \beta \sum_{1}^{K} u_c(q_j)\psi_d(q_i, q_j) = 0$$

  where $\psi_d(q_i, q_j)$ is the discrete analogue of ψ

- Rewriting in Matrix form

$$(\mathbb{I} - \beta\Psi)U_c = 0$$

- The system of linear equation has a nontrivial solution with $U_c >> 0$ if $\frac{1}{\beta}$ is the Eigenvalue of Ψ. Therefore U_c is the Eigenvalue associated with $\beta \in (0, 1)$

- In general, in the discrete case, there could be multiple values of $\beta \in (0, 1)$. Therefore the discrete system is partially identified

Identification Intuition

- In the continuous case, we can define the linear operator A as

$$(Au_c)(q) = \beta \int u_c(q')\psi(q, q')dq'$$

- The Euler Equation implies that

$$u_c = \beta A u_c$$

- Escanciano et al. (2020) shows that if $u_c >> 0$ and $Au_c >> 0$ and A is a compact operator, then a solution for u_c exists if $\beta = \frac{1}{\rho(A)}$, where $\rho(A)$ is the largest real eigenvalue of the operator A.

- Escanciano et al. (2020) shows that there is a unique value for β and u_c that satisfy the Euler equation

Identification of the Joint $\pi(Y, C, R, A)$

- There are two (three?) main challenges in identifying the lifetime consumption distribution from the CPHS data.
  - The data covers only four years
    - Introduce a Markov assumption
  - The data does not contain information on assets and returns
    - Estimating individuals' assets from consumption and income

Problem: Short Panel. Solution: "Learnability"/Markov Assumption

- It is known that we can decompose any joint distribution as

$$\pi(\{Y\}_t^T, \{C\}_t^T, \{R\}_t^T) = \prod \pi(Y_T, C_T, R_T \mid \{Y\}_t^{T-1}, \{C\}_t^{T-1}, \{R\}_t^{T-1})...\pi(Y_0, C_0, R_0)$$

- Unfortunately, we cannot observe each individual's entire life sequence of shocks. Therefore we introduce the following assumptions
  - A1: **stationarity** $\pi_p(\{Y_t\}_{t=0}^T, \{R_t\}_{t=0}^T, \{C_t\}_{t=0}^T, \{V_t\}_{t=0}^T, A_0) = \pi(\{Y_t\}_{t=0}^T, \{R_t\}_{t=0}^T, \{C_t\}_{t=0}^T, \{V_t\}_{t=0}^T, A_0)$ for all $p \in Periods$
  - A2: **Learnablity/M-Markov Process**: There is a (known) M such that for all $t > m$ we

$$\pi(Y_t, C_t \mid Y_{t-1}, C_{t-1}, V_{t-1}, ..., Y_0, C_0, V_0) =$$
$$\pi(Y_t, C_t \mid Y_{t-1}, C_{t-1}, V_{t-1}, ..., Y_{t-M}, C_{t-M}, V_{t-M})$$

- Under this assumption, we can estimate the conditional distribution and stitch them together to identify the uncertainty cost

Problem: Short Panel. Solution: ”Learnability”/Markov Assumption

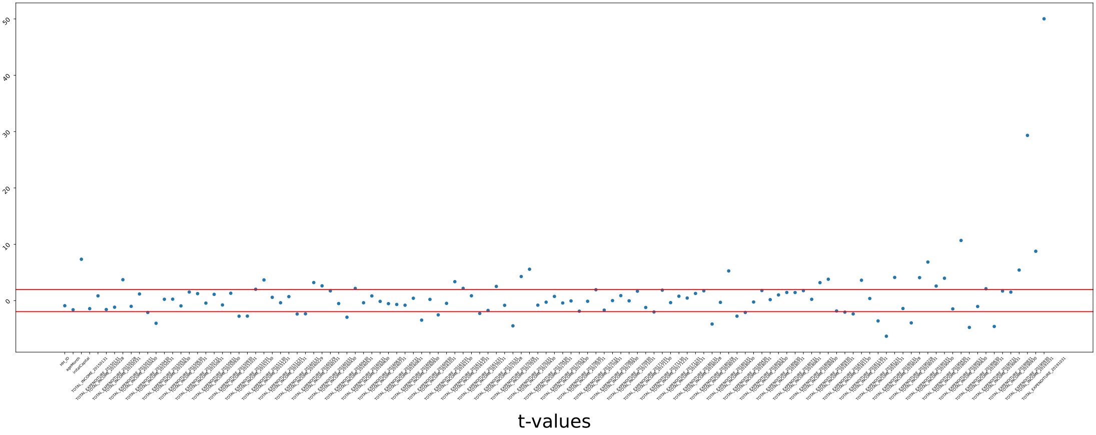


Figure: Consumption - T-Values

Problem: Short Panel. Solution: "Learnability"/Markov Assumption

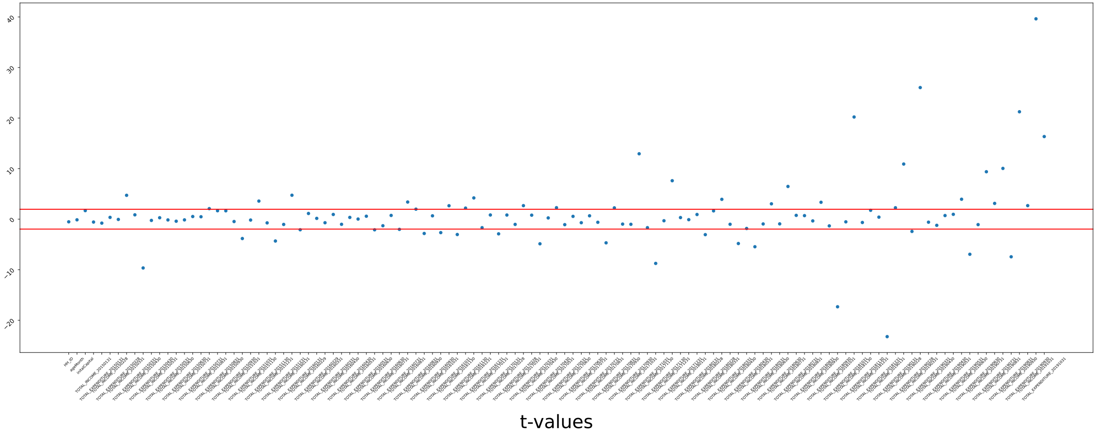
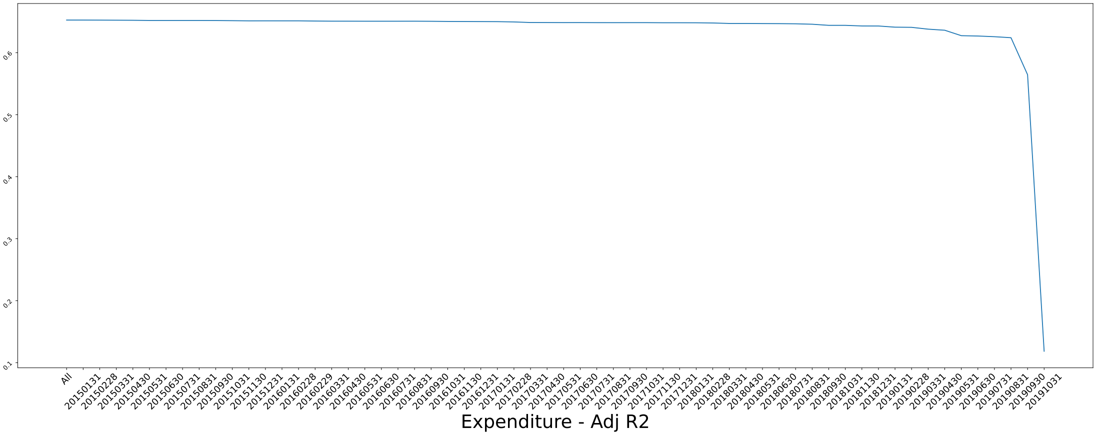
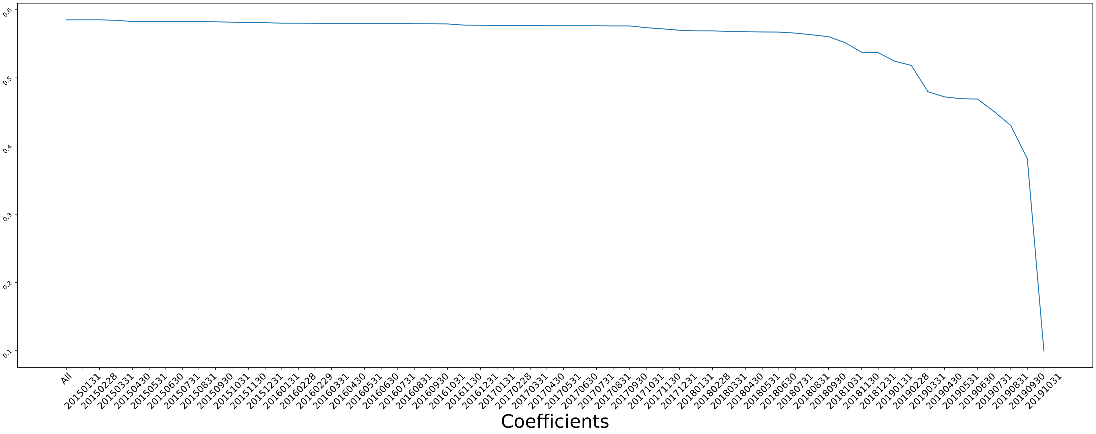


Figure: Income - T-Values

Problem: Short Panel. Solution: "Learnability"/Markov Assumption



Problem: Short Panel. Solution: ”Learnability”/Markov Assumption



Problem: Initial Wealth is not observed. Solution: Estimate it

- We do not observe in the data information on the HH assets
- We do observe information on income from interest (and other similar sources)
- As wealth is usually poorly measured, some literature tried to estimate individuals' wealth from Tax data on non-labor income. (Smith, Zidar and Zwick (2022), Saez and Zucman, (2016), Piketty, Saez and Zucman, (2018))
- I suggest using consumption, labor income, and not-labor income to estimate each HH wealth

Problem: Initial Wealth is not observed. Solution: Estimate it

- We impose the following assumption:
  A3 **(Constant Interest Rate)**: $R_t = R \forall t$

- Let I_{t+1} be the realized returns at period $t+1$ and Notice that

$$I_{t+1} = r_t(A_t + y_t - c_t)$$

  and using the B.C we can derive the following

$$I_{t+1} = r(A_t + y_t - c_t) \implies$$

$$I_{t+1} = r(A_0 + \sum_{\tau=1}^{t}(y_\tau - c_\tau) + \sum_{\tau=1}^{t-1} I_\tau)$$

$$I_{t+1} = r \times A_0 + r \times (\sum_{\tau=1}^{t}(y_\tau - c_\tau) + \sum_{\tau=1}^{t-1} I_\tau)$$

- We can then estimate R and A_0 for each HH using a Fixed Effects regression.

Results - Wealth Distribution

- $R = 0.0051585(0.0003419)$, which implies $R_y \approx 6.3\%$ yearly returns (India 10 years bonds is 6%-7.5%)
- Estimated Wealth Distribution, CPHS (1000 rupees)

(1) mean	(2) sd	(3) min	(4) max	(5) p1	(6) p5	(7) p10	(8) p25	(9) p50	(10) p75	(11) p90	(12) p95	(13) p99
577.3	1,611	-3,217	76,276	-451.9	-164.2	-88.58	-23.50	-0.363	334.7	2,099	3,542	7,412

- NSS AIDIS Wealth Distribution (2013) (1000 rupees)

(1) mean	(2) sd	(3) min	(4) max	(5) p1	(6) p5	(7) p10	(8) p25	(9) p50	(10) p75	(11) p90	(12) p95	(13) p99
1,631	11,277	-84,960	1.850e+06	-117.7	0	25.50	157.3	500	1,384	3,397	5,840	16,468

Estimation of large scale conditional distribution

- We want to be as flexible as we can when estimating the joint distribution
- I suggest using Normalized Flow as a way to estimate the conditional distribution.
- There are three main advantages to this method
  - Allows for fast sampling from the estimated distribution
  - Allows to estimate conditional densities easily
  - We can easily get the density function

Normalized Flows - the idea in a nutshell

- Let $\boldsymbol{x} \sim p(\boldsymbol{x})$ and $u \sim p_u(\boldsymbol{u})$.
- The idea behind Normalized Flows is to express $\boldsymbol{x}$ by a differentiable and invertible transformation of $\boldsymbol{u} \sim p_u(\boldsymbol{u})$

$$x = T(u), u \sim p_u(\boldsymbol{u})$$

- Using this transformation and the change in variables formula, we can express the density of $\boldsymbol{x}$ as

$$p_x(\boldsymbol{x}) = p_u(T^{-1}(\boldsymbol{x}))|detJ_{T^{-1}}(\boldsymbol{x})|^{-1}$$

- where T^{-1} is the inverse of T and J_T is the Jacobian of T
- For continuous variables, and $u \sim U(0,1)$, we know that T exists, as we can use the CDFs on the marginals. For other distributions, we can use an additional transformation of the CDF

Normalized Flows - KL Motivation

- I parameterize $T(\boldsymbol{x})_{\theta}^{-1}$ by θ, and maximizes the implied log-likelihood
- Similar to ML, this would minimize the KL distance

$$
\begin{aligned}
L(\theta) &= D_{KL}(p(^{*}\boldsymbol{x})||p_{\theta}(\boldsymbol{x})) \\
&= C - \mathrm{E}_{x\sim p(x)}[p_{\theta}(\boldsymbol{x})] \\
&= C - \mathrm{E}_{x\sim p(x)}[p_{u}(T_{\theta}^{-1}(x)) + log|detJ_{T_{\theta}^{-1}}(x)|]
\end{aligned}
$$

- Finally, given the transformation $T(\boldsymbol{u})$ I can generate samples from $p_{\theta}(x)$ by drawing from u and inverting T.

Parameterization if T^{-1}

- In practice, for each variable for which I estimate the conditional density, I parameterize T as

$$T^{-1}(x_\tau) = \left(\sum_{W} \Phi(\frac{x_T - \mu_{\mathbf{x}_{-\tau}}}{\sigma_{\mathbf{x}_{-\tau}}}) w_{x_{-\tau}} \right) f(\alpha_{\mathbf{x}_{-\tau}}) + \beta_{\mathbf{x}_{-\tau}}$$

- where each parameter ($\mu_{\mathbf{x}_{-\tau}}, \sigma_{\mathbf{x}_{-\tau}}, \alpha_{\mathbf{x}_{-\tau}}, \beta_{\mathbf{x}_{-\tau}}, w_{\mathbf{x}_{-\tau}}$) is a neural network (X size res blocks) and $f : R \to R_+$. (Similar to Flow++,2019) Losses

Estimation Demonstration

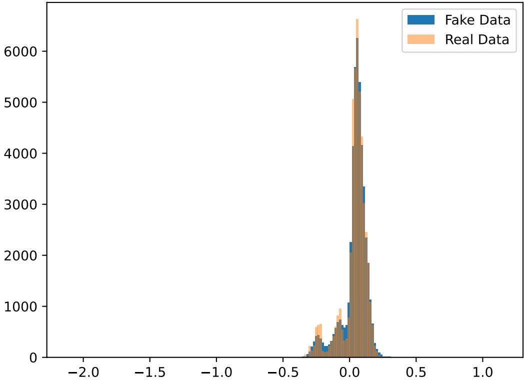


Figure: Income Simulation

Estimation Demonstration

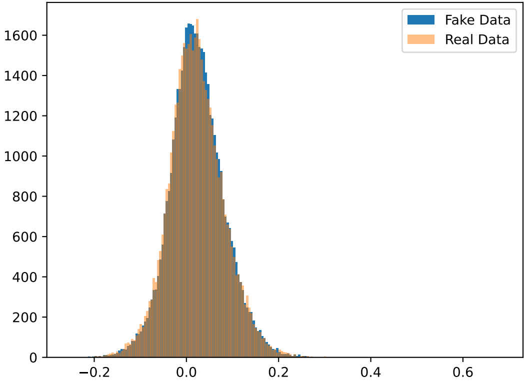
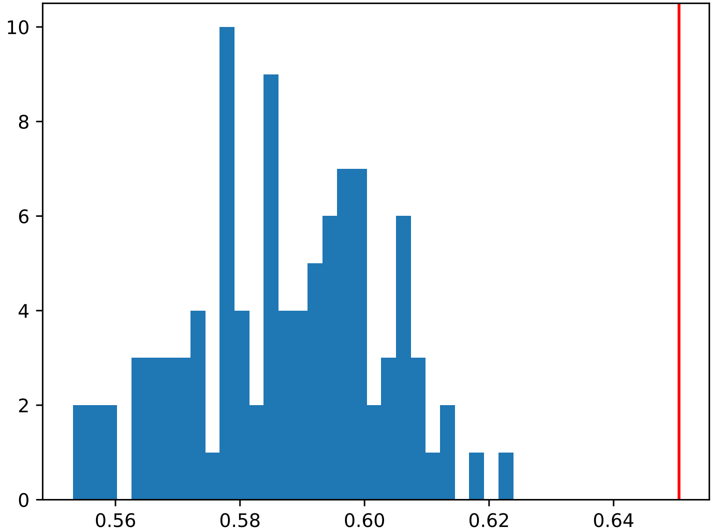
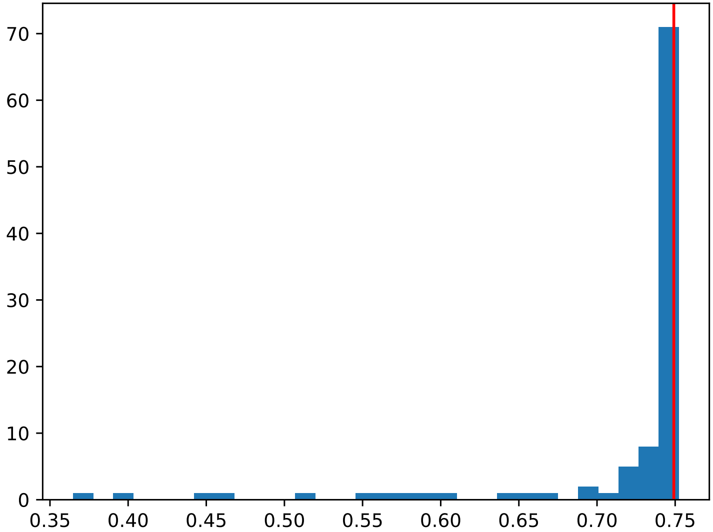


Figure: Consumption Simulation

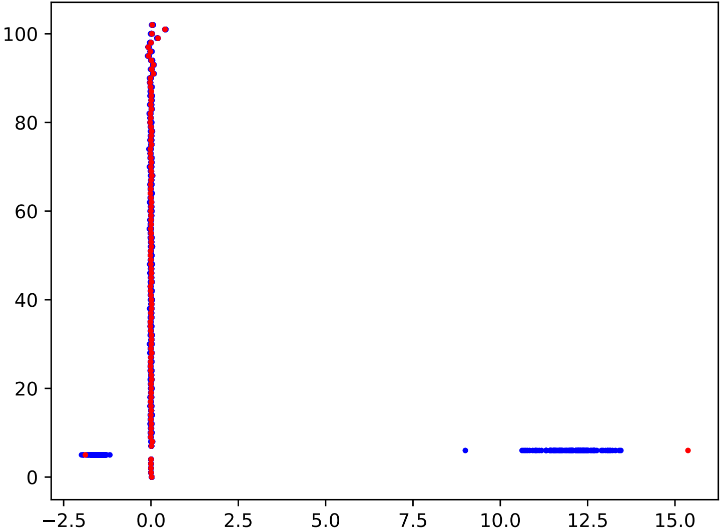
Estimation Demonstration



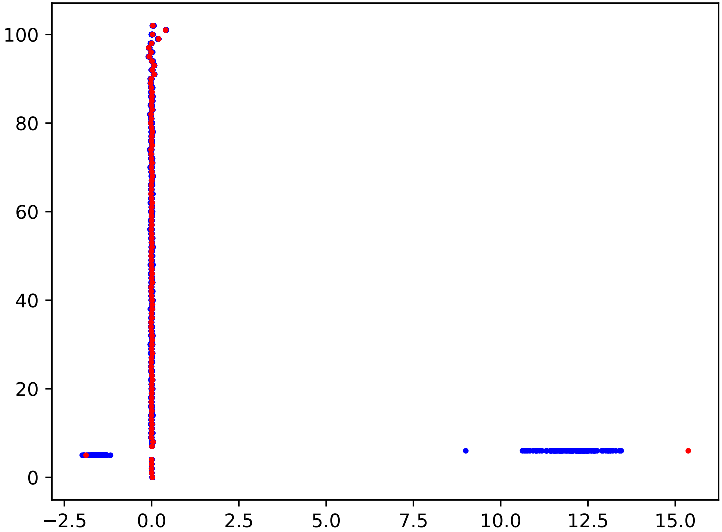
Estimation Demonstration



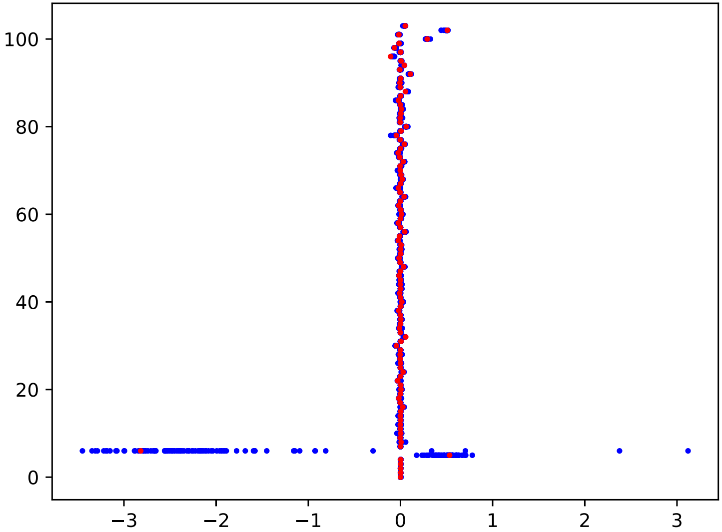
Estimation Demonstration



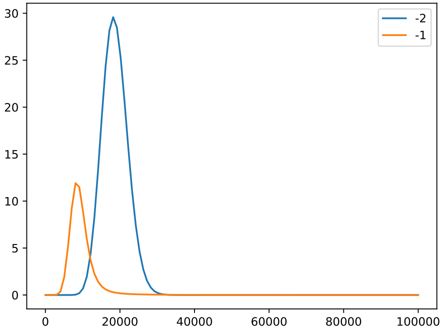
Estimation Demonstration

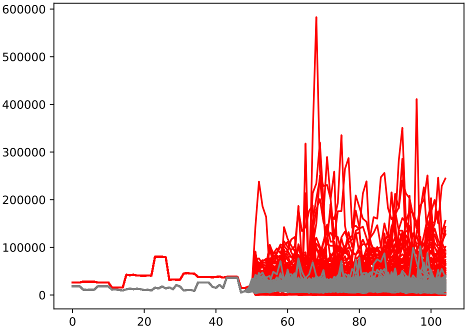


Estimation Demonstration

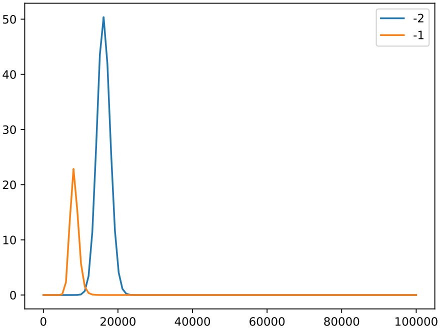


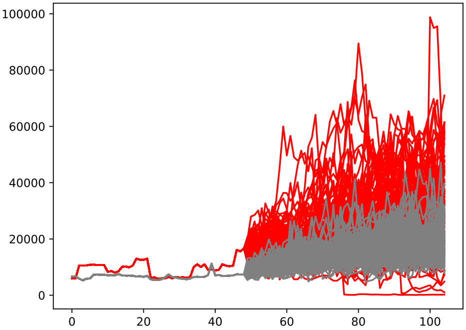
Estimation Demonstration





Estimation Demonstration





Uncertainty Values

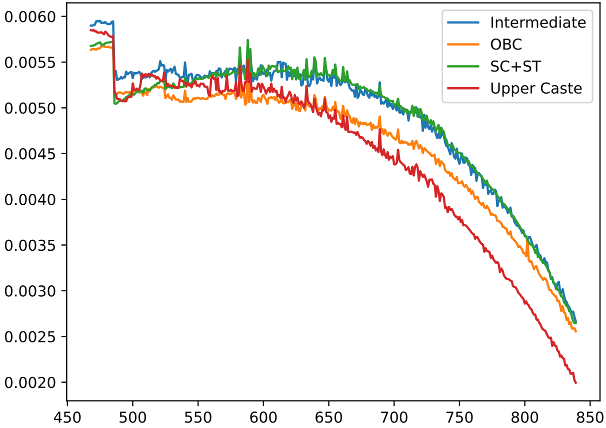


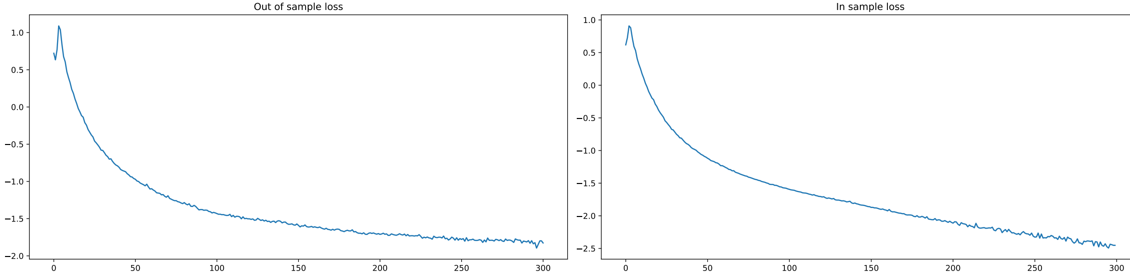
Figure: Cost of uncertainty over the life cycle (In months)

Conclusion

- Utilizing a common utility function across castes, initial findings indicate persistent disparities in the cost of uncertainty.
- These disparities are life-cycle persistent, enduring until age 70.
- In terms of utility, these costs are small, approximate 0.5%, aligning with Lucas (1987).
- The next phase will quantify the proportion of uncertainty disparities attributable to caste-specific preferences.

thx.

Losses


Out of sample loss
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
0
50
100
150
200
250
300
In sample loss
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
0
50
100
150
200
250
300