Estimation of a Latent Reference Point: Method and Application to NYC Taxi Drivers

Marco Stenborg Petterson
Brown University
CSEF - University of Naples Federico II

August 24, 2022

2022 EEA - ESEM

Motivation

- Key assumption in labor supply models: inter-temporal maximization (MaCurdy 1981)
- Consider a taxi driver choosing how many hours to work
  - decision depends on target income
  - target income depends on expectations
- Daily labor supply studies show evidence of income targeting (Camerer et al. 1997)
- Utility of outcomes is experienced relative to some point of reference
  - changes not just levels matters

Motivation

- Reference-dependent preferences have become widely used in economics
  - Labor supply (Camerer et al. 1997, Farber 2008, Farber 2015, Thakral and Tô 2021)
  - Job search (Della Vigna et al. 2017)
  - Consumer choice (Koszegi and Rabin 2006)
  - Housing market (Genesove and Mayer 2001)
- Reference points are **not observed** by the econometrician and **may vary over time**
  - Specific parametric forms of reference point often assumed **a priori**
  - Several theories on formation and evolution of reference points
    - Status quo (Kahneman and Tversky 1979)
    - Forward-looking rational expectations (Koszegi and Rabin 2006)
    - Slow adjusting (Thakral and Tô 2021)

This Paper

- Estimate the evolution of the reference point in a structural model of daily labor supply directly from **observational** data
- Estimate a DDCM with a **latent state** unobserved by econometrician
  - observed and unobserved variables are modeled as a joint Markov process
  - parsimonious model for the transition matrix
- Apply the model to the daily labor supply decisions of NYC taxi drivers
- Tackle open questions on reference point formation
  - how persistent are reference points over time?
  - do agents react differently to positive and negative shocks?(Arkes et al. 2008 and Arkes et al. 2010)

Application

- I consider the daily labor supply decisions of NYC taxi drivers
- I estimate a structural dynamic discrete choice model
  - hours worked and income earned are observed state variables
  - reference point for income is an unobserved state variable
  - utility is time separable and each shift is identical to the next (Camerer et al. 1997)
- 165 million trips, 7 million shifts and 38,659 drivers Data

Structural model: Dynamic programming

- Consider an infinite horizon discrete choice model where V is the agent's lifetime utility from time t onwards

$$V(s_t, x_t, \varepsilon_t) = \max_{a_t} \mathbb{E}\left[\sum_{j=0}^{\infty} \beta^j u(s_{t+j}, x_{t+j}, \varepsilon_{t+j}, a_{t+j}) \,|\, s_t, x_t, \varepsilon_t, a_t\right]$$

- a_t is a binary action, $a_t = 1$ corresponds to ending the shift
- s_t is a state variable observable by both the agent and the econometrician
- x_t is a state variable observable only by the agent
- ε_t is an unobservable distributed as an EV1

Structural model: Dynamic programming

- Consider an infinite horizon discrete choice model where V is the agent's lifetime utility from time t onwards

$$V(s_t, x_t, \varepsilon_t) = \max_{a_t} \mathbb{E}\left[\sum_{j=0}^{\infty} \beta^j u(s_{t+j}, x_{t+j}, \varepsilon_{t+j}, a_{t+j}) \,|\, s_t, x_t, \varepsilon_t, a_t\right]$$

- a_t is a binary action, $a_t = 1$ corresponds to ending the shift
- s_t is a state variable observable by both the agent and the econometrician
- x_t is a state variable observable only by the agent
- ε_t is an unobservable distributed as an EV1

Structural model: Dynamic programming

- Consider an infinite horizon discrete choice model where V is the agent's lifetime utility from time t onwards

$$V\left(s_t, x_t, \varepsilon_t\right) = \max_{a_t} \mathbb{E}\left[\sum_{j=0}^{\infty} \beta^j u\left(s_{t+j}, x_{t+j}, \varepsilon_{t+j}, a_{t+j}\right) \middle| s_t, x_t, \varepsilon_t, a_t\right]$$

- a_t is a binary action, $a_t = 1$ corresponds to ending the shift
- s_t is a state variable observable by both the agent and the econometrician
- x_t is a state variable observable only by the agent
- ε_t is an unobservable distributed as an EV1

Structural model: Dynamic programming

- Consider an infinite horizon discrete choice model where V is the agent's lifetime utility from time t onwards

$$V\left(s_t, x_t, \varepsilon_t\right) = \max_{a_t} \mathbb{E}\left[\sum_{j=0}^{\infty} \beta^j u\left(s_{t+j}, x_{t+j}, \varepsilon_{t+j}, a_{t+j}\right) \middle| s_t, x_t, \varepsilon_t, a_t\right]$$

- a_t is a binary action, $a_t = 1$ corresponds to ending the shift
- s_t is a state variable observable by both the agent and the econometrician
- x_t is a state variable observable only by the agent
- ε_t is an unobservable distributed as an EV1

Structural model: Reference point

- The flow utility is

$$u\left(s_t, x_t, \varepsilon_t, a_t\right) = \begin{cases} u_1\left(s_t, x_t\right) + \varepsilon_{1t} & a_t = 1 \\ \varepsilon_{0t} & a_t = 0 \end{cases}$$

- $s_t = s\left(W_t, H_t\right)$, $x_t = \xi_t$ and

$$u_1\left(W_t, H_t, \xi_t\right) = (1+\alpha)\left(W_t - g(W_t, H_t, \xi_t)\right) - \frac{\psi}{1+\eta}\left(H_t\right)^{1+\eta} \quad W_t < g(W_t, H_t, \xi_t)$$

$$u_1\left(W_t, H_t, \xi_t\right) = (1-\alpha)\left(W_t - g(W_t, H_t, \xi_t)\right) - \frac{\psi}{1+\eta}\left(H_t\right)^{1+\eta} \quad W_t \geq g(W_t, H_t, \xi_t)$$

- $\alpha > 0$ controls the change in marginal utility at the reference point $g\left(W_t, H_t, \xi_t\right)$

α

Structural model: Reference point

- The flow utility is

$$u\left(W_t, H_t, \xi_t, \varepsilon_t, a_t\right) = \begin{cases} u_1\left(W_t, H_t, \xi_t\right) + \varepsilon_{1t} & a_t = 1 \\ \varepsilon_{0t} & a_t = 0 \end{cases}$$

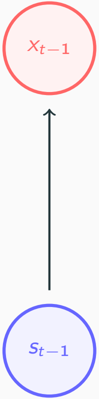
- The reference point is

$$g(W_t, H_t, \xi_t) = f\left(W_t, H_t\right) + \xi_t$$

- Several possibilities for f
  - $f\left(W_t, H_t\right) = 0 \; \forall W_t, H_t$
  - $f\left(W_t, H_t\right)$ is a function of the expected end of shift income at $s_t = \left(W_t, H_t\right)$
- If both s and x were observable, the model would mirror the assumptions and setup of Rust (1987)

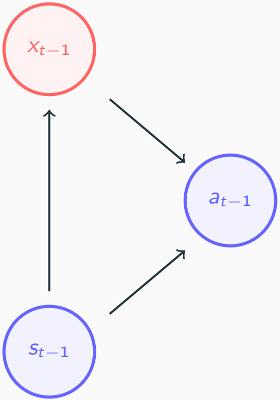
Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



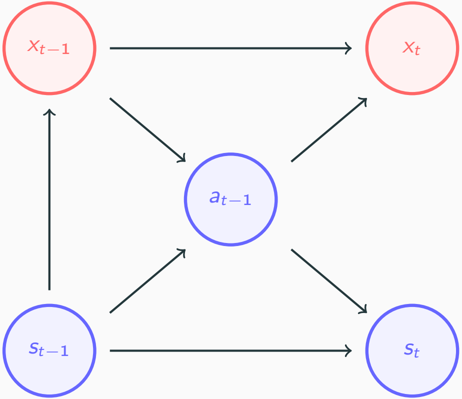
Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



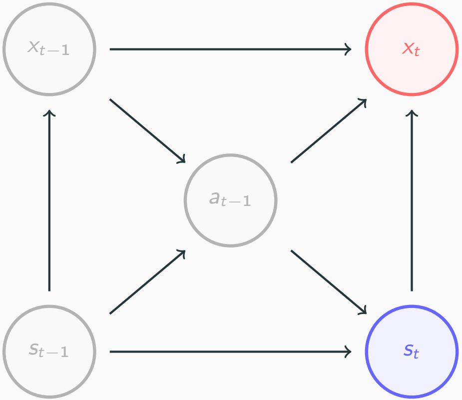
Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



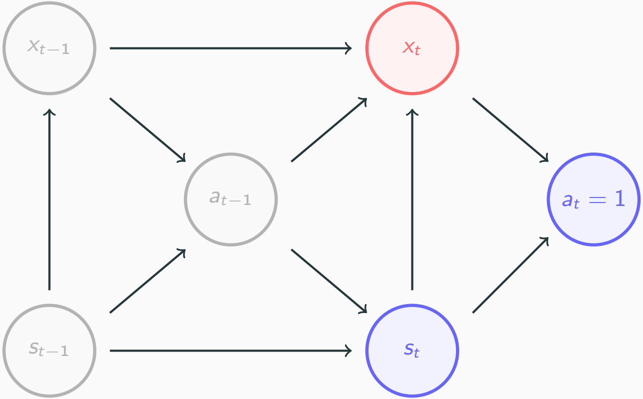
Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



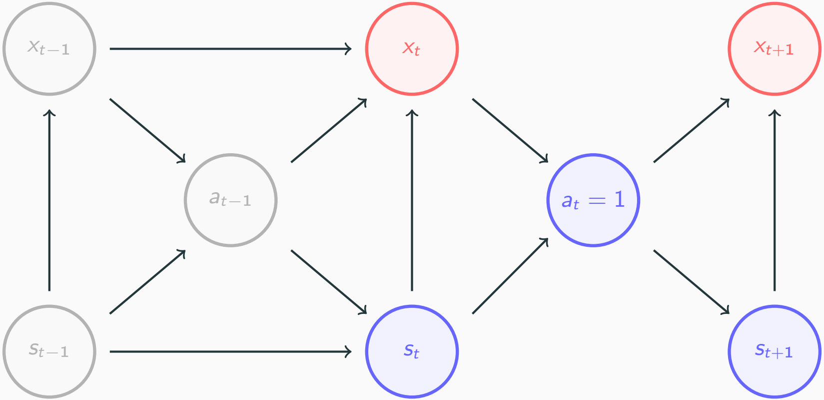
Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



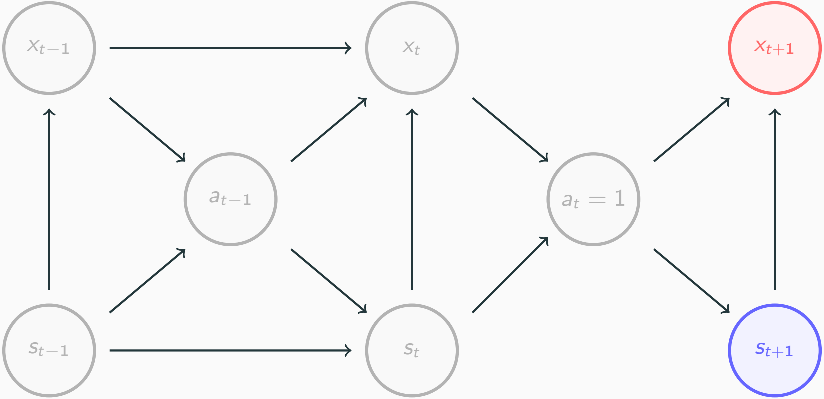
Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



Structural model: Transition matrix

- The states s_t and x_t behave as a joint Markov process



Structural model: Transition matrix

- Formally, the transition probabilities can be stated as

$$\Pr\left(s_{t+1}=j|s_t=k,a_t=0,x_t,x_{t+1}\right)=\Pr\left(s_{t+1}=j|s_t=k,a_t=0\right)=\begin{cases}\theta^S_{jk} & \text{if } j\geq k\\ 0 & \text{otherwise}\end{cases}$$

$$\begin{aligned}\Pr\left(x_{t+1}=m|s_{t+1}=j,s_t=k,x_t=q,a_t=0\right)&=\Pr\left(x_{t+1}=m|s_{t+1}=j,x_t=q,a_t=0\right)\\&=\theta^X_{mqj}\end{aligned}$$

- If $a_t=1$ then $x_{t+1}=1$ and $s_{t+1}=1$ Transition Matrix Example

Parametric transition matrix

- For estimation I adopt a parametric structure for the evolution of the unobserved state

$$
\begin{aligned}
\Pr\left(x_t, s_t | s_{t-1}, x_{t-1}, a_{t-1} = 0\right) = \\
= \Pr\left(x_t | s_t, x_{t-1}, a_{t-1} = 0\right) \Pr\left(s_t | s_{t-1}, a_{t-1} = 0\right) \\
= \frac{\exp\left(h_X\left(x_t, x_{t-1}\right) + h_S\left(x_t, s_t\right)\right)}{\sum_{x'} \exp\left(h_X\left(x'_t, x_{t-1}\right) + h_S\left(x'_t, s_t\right)\right)} \Pr\left(s_t | s_{t-1}, a_{t-1} = 0\right)
\end{aligned}
$$

Functions

- Reference point transition probability depends on:
  - past reference point
  - current realization of observables

Likelihood

- The likelihood is

$$L_N\left(\theta^X, \theta^S, \alpha, \psi, \eta\right) = \prod_{i=1}^{N} \int p\left(a_i, s_i, x_i | \theta^X, \theta^S, \alpha, \psi, \eta\right) d\left(x_i\right)$$

where $a_i = \{a_{i,t}\}_{t=1}^{T_i}$,$s_i = \{s_{i,t}\}_{t=1}^{T_i}$,$x_i = \{x_{i,t}\}_{t=1}^{T_i}$

$$p\left(a_i, s_i, x_i | \theta^X, \theta^S, \alpha, \psi, \eta\right) = \prod_{t=1}^{T_i} p\left(a_{i,t} | a_{i,t-1}, s_{i,t}, x_{i,t}; \theta^X, \theta^S, \alpha, \psi, \eta\right)$$
$$q\left(s_{i,t}, x_{i,t} | s_{i,t-1}, x_{i,t-1}, a_{i,t-1}; \theta^X, \theta^S, \alpha, \psi, \eta\right)$$

Particle Filter

- In order to calculate the integral I resort to particle filter

$$\int p\left(a_i, s_i, x_i | \theta^X, \theta^S, \alpha, \psi, \eta\right) d\left(x_i\right)$$
$$\approx \prod_{t=1}^{T_i} \frac{1}{M} \sum_{m=1}^{M} p\left(a_{i,t} | a_{i,t-1}, s_{i,t}, x_{i,t}^m\right) q\left(s_{i,t}, x_{i,t}^m | s_{i,t-1}, x_{i,t-1}^m, a_{i,t-1}; \theta^X, \theta^S, \alpha, \psi, \eta\right)$$

- Since x_t is discrete we can use a discrete filter taken from the Hidden Markov Models literature
  - overcomes difficulties in choice of number of particles M Discrete PF

Estimation

- I take a fully Bayesian approach to estimation
  - priors are (truncated) normals centered at 0 Priors
  - Flury and Shephard (2011) unbiased likelihood approximations inside Metropolis-Hastings results in exact posterior
- Given the length of the sample for several individuals I estimate the model separately for each individual
  - smaller state space
  - characterizes heterogeneity within the sample

Estimation Procedure

- Initialize $y = \{\alpha, \eta, \psi, \gamma_1, \ldots, \gamma_6\}$
- For $j = 1 : N$
  - draw a candidate $y^* \sim q\left(\cdot|y^{j-1}\right)$
  - solve DP for y^* with VFI and Newton-Kantorovich
  - calculate likelihood with particle filtering
  - set $y^j = y^*$ with probability

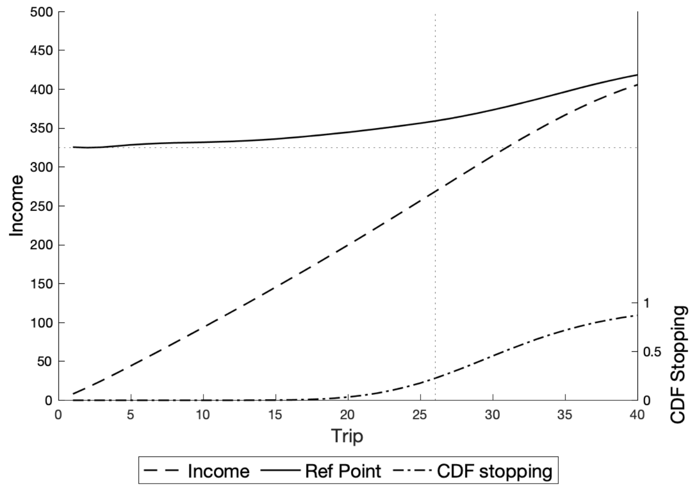
$$\mu = \min\left\{1, \frac{p\left(y^*\right)/q\left(y^*|y^{j-1}\right)}{p\left(y^{j-1}\right)/q\left(y^{j-1}|y^*\right)}\right\}$$

otherwise $y^j = y^{j-1}$

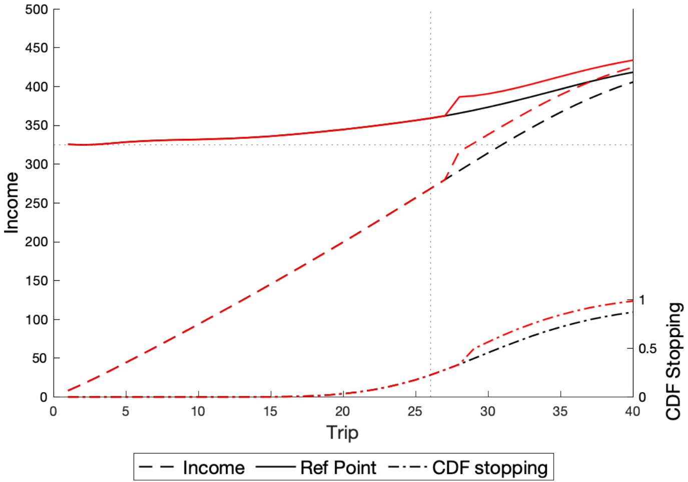
Impulse Response Functions

- I consider an individual chosen at random among those with more than $5,000$ trips
  Distribution Observations Estimates
- I simulate $100,000$ shifts and plot
  - the average reference point
  - the average earned income at every trip
  - the cumulative stopping probability
- I then introduce a shock in the earned income at different times of the shift
  - study the evolution of the reference point and stopping probability after the shock
- I tackle two open questions in the literature
  - how persistent are reference points over time? Persistence
  - do agent react differently to positive and negative shocks? Asymmetric response

Impulse Response Functions CDF Hours Plot



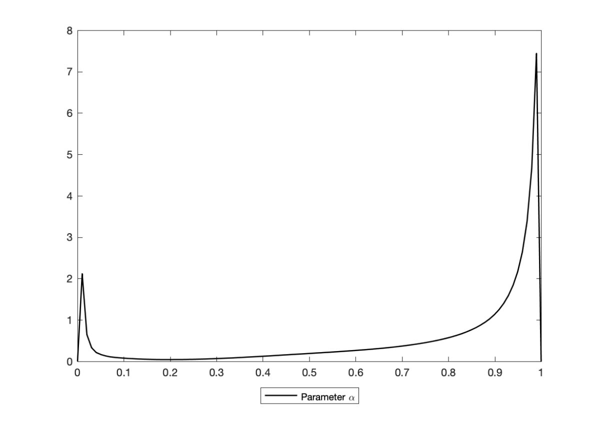
Impulse Response Functions CDF Hours Plot



Plot of MAP in the sample

- We consider individual with more than $5,000$ trips and plot the kernel density of the Maximum a Posteriori for some parameters Utility MAP

α: change in marginal utility



Conclusions

- It is possible to identify patterns in the formation and evolution of reference points using **observational** data
  - more generally applicable to models with hidden states
- It is possible to **test** theories on reference point formation
- Future research:
  - test in experimental setting, other applications (e.g. consumer choice)
  - strategic interaction: evidence of reference-dependent preferences in sequential bargaining (coming soon!)

Appendix

Observable Transition Matrix

Suppose $a_t \in \{0,1\}$, $S_t \in \{A, B\}$ and $X_t \in \{a, b\}$

	a_0, S_A	a_0, S_B
a_0, S_A	$P_{Aa}\theta^S_{AA}\theta^X_{aaA}P^X_{aA} + P_{Ab}\theta^S_{AA}\theta^X_{baA}P^X_{aA}$ $+P_{Aa}\theta^S_{AA}\theta^X_{abA}P^X_{bA} + P_{Ab}\theta^S_{AA}\theta^X_{bbA}P^X_{bA}$	$P_{Ba}\theta^S_{BA}\theta^X_{aaA}P^X_{aA} + P_{Bb}\theta^S_{BA}\theta^X_{aaA}P^X_{aA}$ $+P_{Ba}\theta^S_{BA}\theta^X_{abA}P^X_{bA} + P_{Bb}\theta^S_{BA}\theta^X_{abA}P^X_{aA}$
a_0, S_B	0	$P_{Ba}\theta^S_{BA}\theta^X_{aaA}P^X_{aB} + P_{Bb}\theta^S_{BA}\theta^X_{aaA}P^X_{aB}$ $+P_{Ba}\theta^S_{BA}\theta^X_{abA}P^X_{bB} + P_{Bb}\theta^S_{BA}\theta^X_{abA}P^X_{aB}$

- Where
  - p_{Aa} is the CCP of $a_t = 0$ if $x_t = a$ and $s_t = A$
  - θ^S_{AB} is $\Pr(s_t = A | s_{t-1} = B, a_{t-1} = 0)$
  - $\theta^X_{abA}$is $\Pr(x_t = a | x_{t-1} = b, s_{t-1} = A, a_{t-1} = 0)$
  - P^X_{aA} is $\Pr(x_{t-1} = a | s_{t-1} = A, a_{t-1} = 0)$

Complete Transition Matrix

	a_0, S_A, X_a	a_0, S_A, X_b	a_0, S_B, X_a	a_0, S_B, X_b
a_0, S_A, X_a	$P_{Aa}\theta^S_{AA}\theta^X_{aaA}$	$P_{Ab}\theta^S_{AA}\theta^X_{baA}$	$P_{Ba}\theta^S_{BA}\theta^X_{aaA}$	$P_{Bb}\theta^S_{BA}\theta^X_{aaA}$
a_0, S_A, X_b	$P_{Aa}\theta^S_{AA}\theta^X_{abA}$	$P_{Ab}\theta^S_{AA}\theta^X_{bbA}$	$P_{Ba}\theta^S_{BA}\theta^X_{abA}$	$P_{Bb}\theta^S_{BA}\theta^X_{abA}$
a_0, S_B, X_a	0	0	$P_{Ba}\theta^S_{BB}\theta^X_{aaB}$	$P_{Bb}\theta^S_{BB}\theta^X_{aaB}$
a_0, S_B, X_b	0	0	$P_{Ba}\theta^S_{BB}\theta^X_{abB}$	$P_{Bb}\theta^S_{BB}\theta^X_{abB}$

Back Structure Back Identification

Parametric Transition

$$h_X = \gamma_1 \mathbb{1}\{x_t - x_{t-1} < 0\}|x_t - x_{t-1}| + \gamma_2 \mathbb{1}\{x_t - x_{t-1} > 0\}|x_t - x_{t-1}|$$

$$h_S =$$

$$\mathbb{1}\{f(s_t) \geq f(s_1)\}\left(\gamma_3 \mathbb{1}\{g(s_t, x_t) - f(s_t) < 0\}|g(s_t, x_t) - f(s_t)| + \gamma_4 \mathbb{1}\{g(s_t, x_t) - f(s_t) > 0\}|g(s_t, x_t) - f(s_t)|\right) +$$

$$\mathbb{1}\{f(s_t) < f(s_1)\}\left(\gamma_5 \mathbb{1}\{g(s_t, x_t) - f(s_t) < 0\}|g(s_t, x_t) - f(s_t)| + \gamma_6 \mathbb{1}\{g(s_t, x_t) - f(s_t) > 0\}|g(s_t, x_t) - f(s_t)|\right)$$

Back MAP

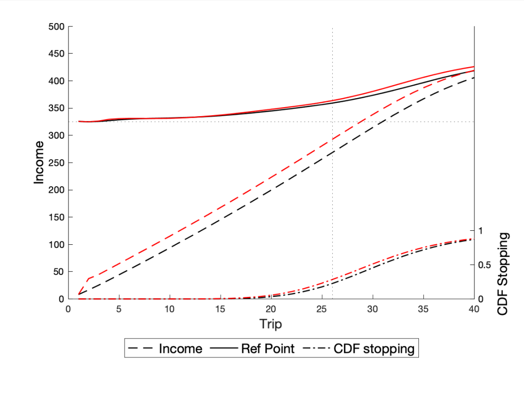
Priors

- $\psi \sim \mathcal{N}(0, 10)\,\mathbb{1}\{\psi > 0\}$
- $\eta \sim \mathcal{N}(0, 10)\,\mathbb{1}\{\eta > 0\}$
- $\alpha \sim \mathcal{N}(0, 10)\,\mathbb{1}\{\alpha \geq 0\}$
- $\gamma_1, \ldots, \gamma_6 \sim \mathcal{N}(0, 10)$ Back

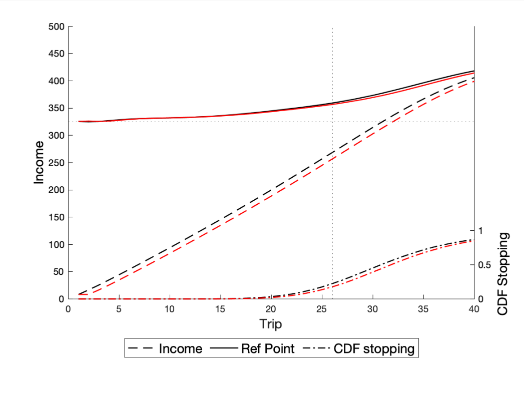
Impulse Response Functions: Early Shock Back

- The reference point is not very persistent, the impact of the shock to observable disappears after few trips

Positive Shock



Negative Shock



Fixed Point

- Consider an example where $a \in \{0,1\}$, $s \in \{A,B,C,D\}$ and $x \in \{\alpha,\beta,\gamma\}$
- Write equation the system of equations as

$$\begin{aligned}\Pr(a_t{=}0|s_t{=}A,s_{t-1}{=}A,a_{t-1}{=}0) = & P_{A\alpha}\big(\theta^X_{\alpha\alpha A}\Gamma_{A\alpha}+\theta^X_{\alpha\beta A}\Gamma_{A\beta}+\theta^X_{\alpha\gamma A}\Gamma_{A\gamma}\big)+ \\ & +P_{A\beta}\big(\theta^X_{\beta\alpha A}\Gamma_{A\alpha}+\theta^X_{\beta\beta A}\Gamma_{A\beta}+\theta^X_{\beta\gamma A}\Gamma_{A\gamma}\big)+P_{A\gamma}\big(\theta^X_{\gamma\alpha A}\Gamma_{A\alpha}+\theta^X_{\gamma\beta A}\Gamma_{A\beta}+\theta^X_{\gamma\gamma A}\Gamma_{A\gamma}\big)\end{aligned}$$

which can in turn be rewritten as

$$\begin{aligned}P^{00}_{AA}-P_{A\gamma} = & \big(P_{A\alpha}-P_{A\gamma}\big)\big(\theta^X_{\alpha\alpha A}\Gamma_{A\alpha}+\theta^X_{\alpha\beta A}\Gamma_{A\beta}+\theta^X_{\alpha\gamma A}\big(1-\Gamma_{A\alpha}-\Gamma_{A\beta}\big)\big) \\ & +\big(P_{A\beta}-P_{A\gamma}\big)\big(\theta^X_{\beta\alpha A}\Gamma_{A\alpha}+\theta^X_{\beta\beta A}\Gamma_{A\beta}+\theta^X_{\beta\gamma A}\big(1-\Gamma_{A\alpha}-\Gamma_{A\beta}\big)\big)\end{aligned}$$

where $\Gamma_{kq} = \frac{\Pr(a_{t-1}=0,s_{t-1}=k,x_{t-1}=q)}{\sum_g \Pr(a_{t-1}=0,s_{t-1}=k,x_{t-1}=g)}$

Fixed Point

Proposition *Let F be a fixed point operator associated with the system represented by equation 1 and denote by J its Jacobian with respect to θ^X. If the highest eigenvalue of Jacobian matrix is smaller than 1 then the system of equations has a unique solution, and the estimator for θ^X is identified.*

- The result follows from Banach's contraction theorem Back

Discrete Particle Filter

$$\text{initialization} \begin{cases} \tilde{\pi}_1 & = \pi_1 = \Pr\left(a_0, s_0, x_0\right) \\ \log \rho_1 & = 0 \end{cases}$$

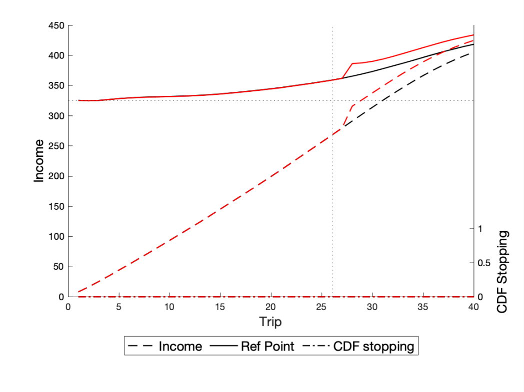
$$\text{iteration} \begin{cases} \pi_{t+1} & = \tilde{\pi}_t Q_{t+1} \\ \tilde{\pi}_{t+1} & = \frac{\pi_{t+1}}{\|\pi_{t+1}\|_1} \\ \log \rho_{t+1} & = \log \rho_t + \|\pi_{t+1}\|_1 \end{cases}$$

where $Q_{t+1,xx'} = \mathbb{P}\left(x_{t+1} = x', s_{t+1}, a_{t+1} \mid x_t = x, s_{t+1}, a_{t+1}\right)$ i.e. the transition matrix of the unobserved state variable keeping fixed the observed state variables at the observed values. Back

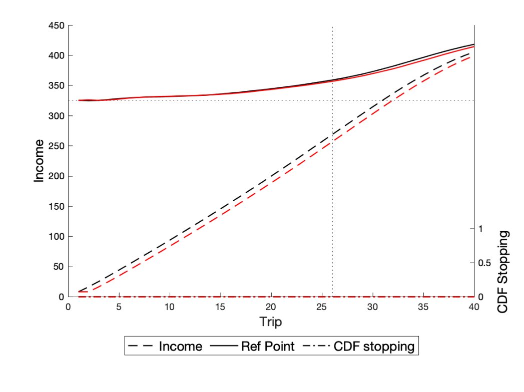
Selection

- We can account for selection by forcing individuals to keep working Back

Positive Shock



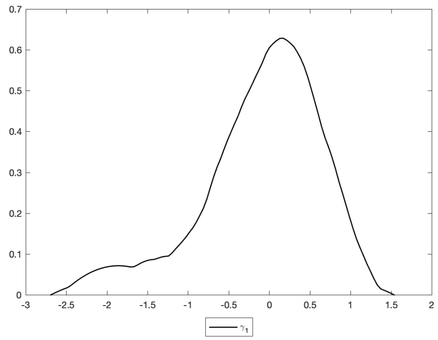
Negative Shock



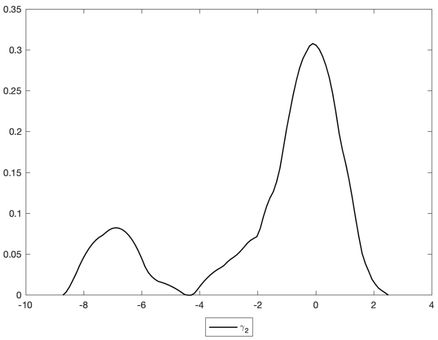
MAP: Change in reference point

Transition Back

γ_1: Negative path dependence



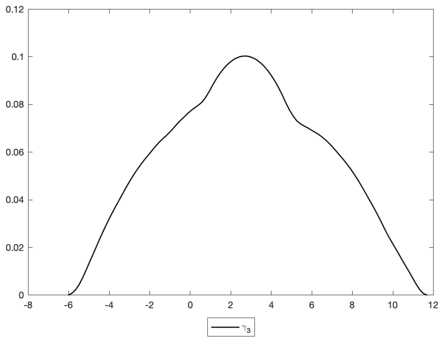
γ_2: Positive path dependence



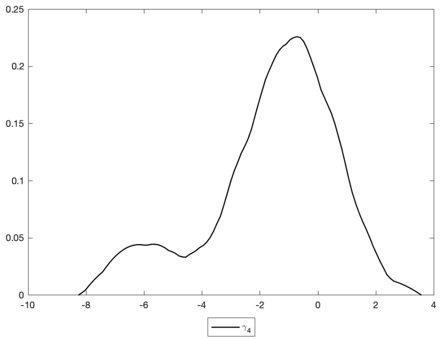
MAP: Change observable lower than average

Transition Back

γ_3: Negative change



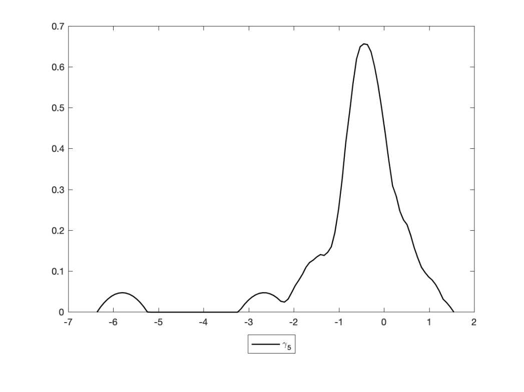
γ_4: Positive change



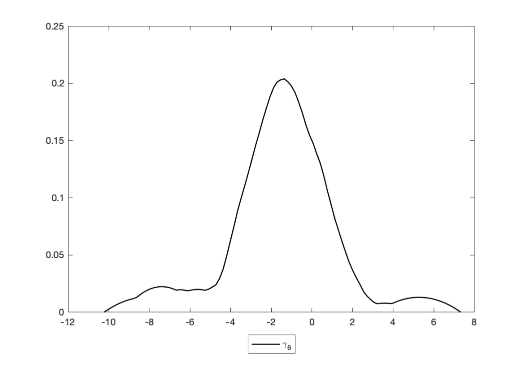
MAP: Change observable higher than average

Transition Back

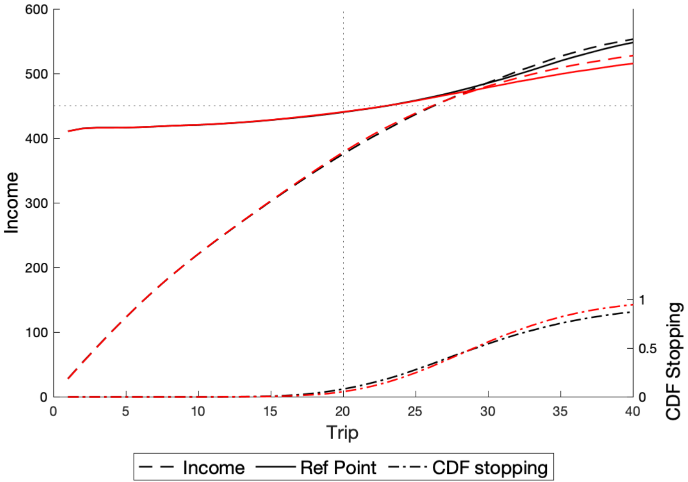
γ_5: Negative change



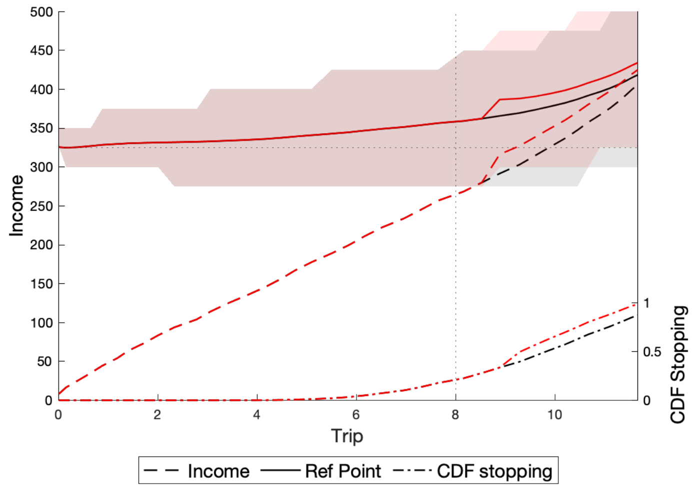
γ_6: Positive change



Reference Points Matter Back



Impulse Response with Hours Back



Identification

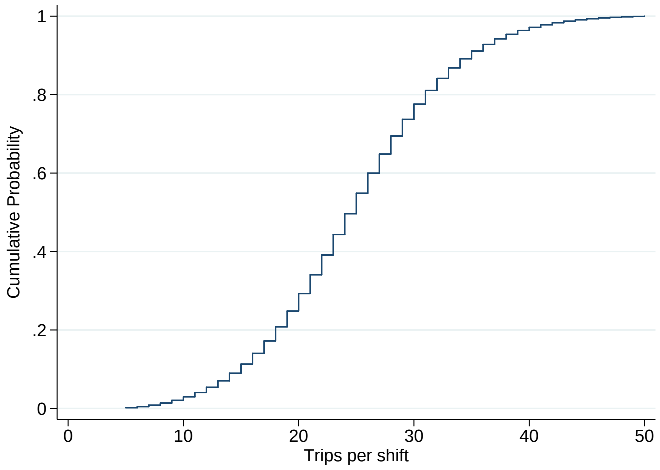
- Follows Connault (2016)
- Relies on the following system of equations having a unique solution

$$\begin{aligned} & \Pr(a_t{=}0|,s{=}j,a_{t-1}{=}0,s_{t-1}{=}k) \\ & {=}\sum_m \Pr(a_t{=}0|s_t{=}j,x_t{=}m,a_{t-1}{=}0,s_{t-1}{=}k)\Pr(x_t{=}m|s_t{=}j,a_{t-1}{=}0,s_{t-1}{=}k) \\ & {=}\sum_m P_{jm}\sum_q\left[\theta^X_{mqj}\frac{\Pr\left(a_{t-1}{=}0,s_{t-1}{=}k,x_{t-1}{=}q\right)}{\sum_g \Pr\left(a_{t-1}{=}0,s_{t-1}{=}k,x_{t-1}{=}g\right)}\right] \end{aligned} \tag{1}$$

where θ^S_{jk} is observable and P_{jm} is the Conditional Choice Probability which is a function of θ^S and θ^X Transition Matrix Example Fixed Point

- Identification relies on the unobserved state x influencing the CCPs, and CCPs depending in turn on
  - utility parameters
  - both observable and unobservable components of the transition matrix Back

Distribution of trips IR

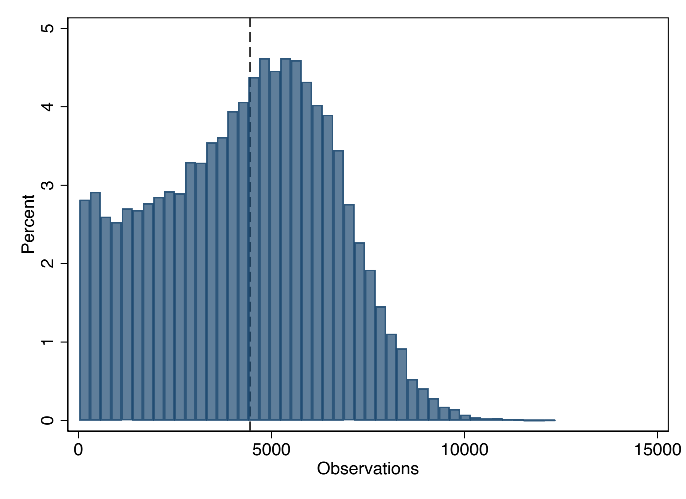


Data Back

- NYC Taxi and Limousine Commission trip sheet data for 2013
  - earnings
  - start and end times of each trip
- Shift: **consecutive trips** of the same driver with less than 6 hours break
- Observed variables are cumulative earnings and hours worked within each shift
  - Cumulative income is discretized at interval of 25$ while cumulative hours worked at 1 hour intervals
- 165 million trips, 7 million shifts and $38,659$ drivers

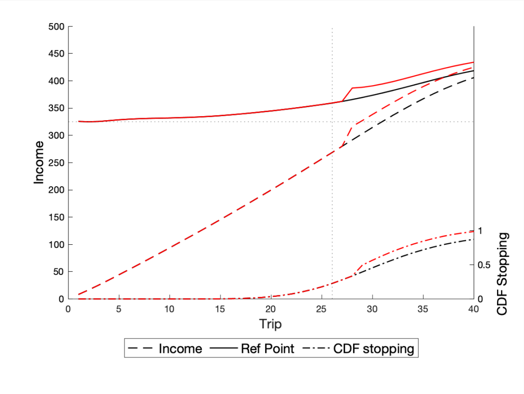
Summary Statistics: Observation per individual

IR

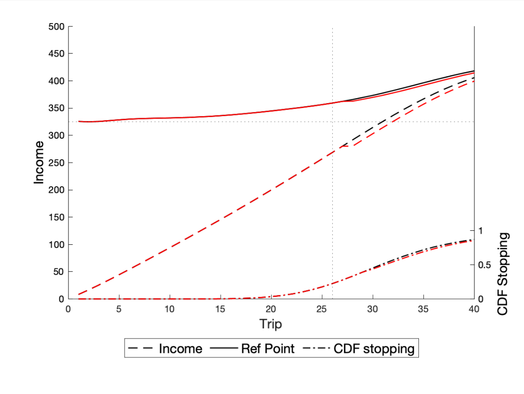


- The reference point reacts much more to a positive shock in observable income than a negative one

Positive Shock



Negative Shock



Estimates IR

- Estimates for individual

Parameters	
ψ	0.425
	[0.021,0.581]
η	0.0912
	[0.017,1.608]
α	0.956
	[0.949,0.961]
γ_1	-0.566
	[-0.668,-0.523]
γ_2	-2.11
	[-2.242,-2.042]
γ_3	0.364
	[0.327,0.467]
γ_4	1.974
	[1.904 ,2.103]
γ_5	-1.97
	[-2.058 ,0.177]
γ_6	-0.79
	[-1.13 ,-0.541]