Harold Zurcher and Paul Volcker— The Minutes of a Meeting That Never Was

Philipp Müller
University of Zurich

Gregor Reich
Tsumcor Research AG

January 24, 2022

Harold Zurcher
(1926—2020)

Maintenance manager at
Madison (Wisconsin) Metropolitan Bus Company
from 1955 to 1993

Paul Volcker
(1927—2019)

Chairman of the
United States Federal Reserve System
from 1979 to 1987

Outline

- The discount factor in DDCMs is considered to be hard to estimate due to its alleged **"poor identification"** (Rust, 1987; Aguirregabiria and Mira, 2010, and others)
- We propose a new method to *systematically* address this problem using **homotopy path continuation**
- In the "bus engine replacement model" of Rust (1987), we find—against common belief—the discount factor to be **well identified**, and *significantly* **larger than 1**
- Since discount factors greater than one are *unpopular*, we use a **natural experiment** to validate this finding:
  - Volcker's appointment curbed inflation (and negative *real* interest rates), happening in the middle of our sample
  - We add a **structural break** in Zurcher's discounting, finding that the *change* of the discount factor itself as well as the estimated *timing* strongly support our estimate

Harold Zurcher
(1926—2020)

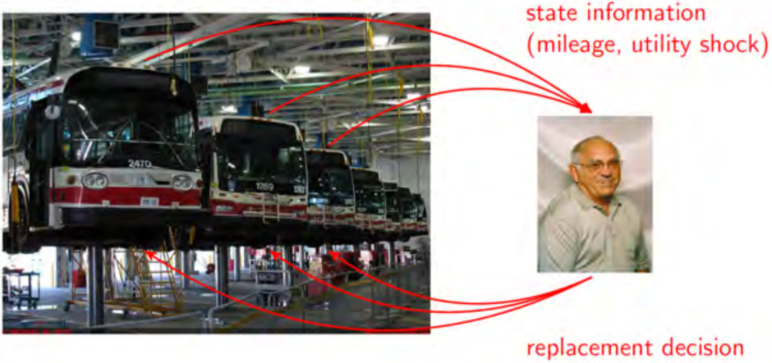
Maintenance manager at
Madison (Wisconsin)
Metropolitan Bus Company
from 1955 to 1993

Paul Volcker
(1927—2019)

Chairman of the
United States
Federal Reserve System
from 1979 to 1987

The Bus Engine Replacement Model (Rust, 1987)

John Rust: *Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher.* Econometrica, 1987.



Rust (1987) pioneered the estimation of dynamic discrete choice models, with decisions based on dynamic programming (Bellman)

Harold Zurcher's Discount Factor Puzzle

- Estimating Zurcher's valuation of the future
  - To value and compare payoff streams in the future, Zurcher must discount them to a common date. This is formalized by the discount factor, here: β.
  - Under certain conditions, β in DDCMs can be estimated
- Rust (1987) attempted to do so, but failed:

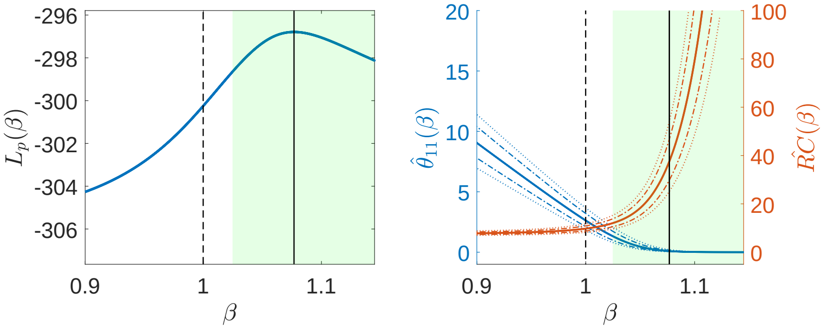
  > *I was not able to precisely estimate the discount factor [...] if I treated β as a free parameter, the estimated information matrix was nearly singular, causing difficulties for the maximization algorithm.*

  However, he adds:

  > *I did note a systematic tendency for the estimated value of β to be driven to 1. This curious behavior may be an artifact of computer round-off errors, or it could indicate a deeper result.*

A Resolution of the Puzzle?

- We address this problem using
  - our homotopy path continuation estimation approach
  - relative value iteration to account for divergence as $\beta \to 1$
- We find the estimate for β
  - to be clearly identified
  - to be significantly larger than 1



Estimation Results and Robustness

	Rust (1987)	Müller and Reich		
p	$\{1,2,3,4\}$	$\{1,2,3,4\}$	$\{1,2,3\}$ $\{4\}$	$\{1,2,3\}$ $\{4,5,7\}$ $\{6,8\}$
$\hat{\beta}$	0.9999 —	1.0768 $[1.0245, \infty)$	1.0467 [0.9897, 1.1042]	1.0283 [1.0073, 1.0476]
LL	-6,055.25	-6,051.79	-6,011.51	-11,097.82
p-value ($H_0 : \beta = 1$)	—	0.0086	0.1552	0.0092

On the "Acceptable Domain" of Discount Factors

- An editor:

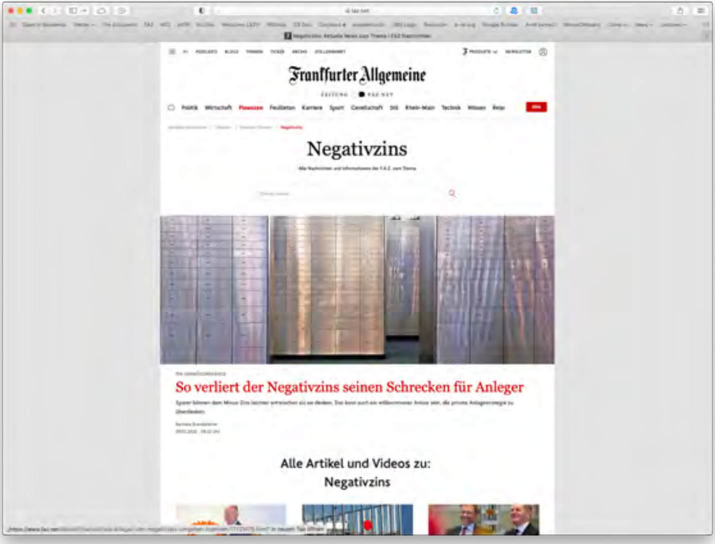
  Discount factors larger than one make little economic sense. I would interpret your results in that application as evidence for misspecification.

- Others accept $\beta \geq 1$ on various grounds:
  - Behavioral (Erdem and Keane, 1996):

    The intertemporal factor is usually assumed to be between 0 and 1 because it is assumed to be 1 / (1 + interest rate) although behaviorally this does not have to hold.

  - Infinite horizon dynamic optimization problems can be viewed approximations of finite horizon problems (Morton and Wecker, 1977), justifying relative value iteration
  - Kocherlakota (1990) theoretically treats discount factors greater than one in macro-economic growth models; Heaton (1995) estimates them
  - Economic definition: **Negative interest rates** imply discount factors greater than one

Negative Interest Rates are Real



Harold Zurcher
(1926—2020)

Maintenance manager at
Madison (Wisconsin)
Metropolitan Bus Company
from 1955 to 1993

Paul Volcker
(1927—2019)

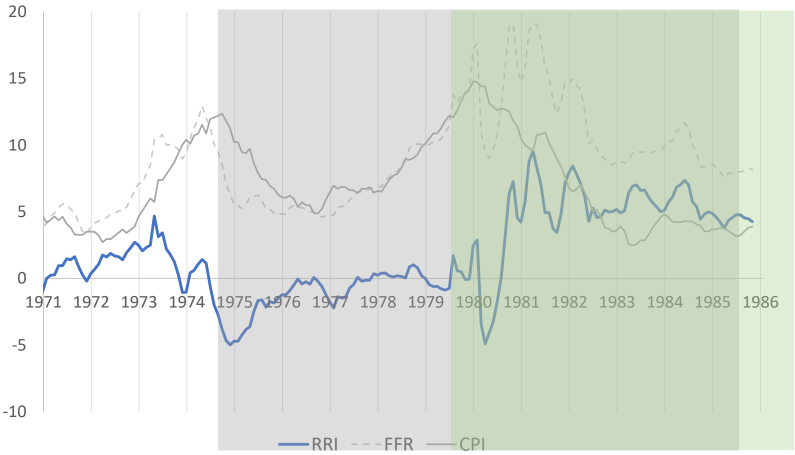
Chairman of the
United States
Federal Reserve System
from 1979 to 1987

Paul Volcker

- In July 1979, upon fears of a recession and inflation of about 12% in the United States, President Jimmy Carter nominates Paul Volcker as new Chairman of the Federal Reserve
- In the Congressional hearning, Volcker testifies:

  > *the American people have, I suspect, become convinced as never before that inflation is here to stay and that it may rise. That affects activity; it affects the way they invest; it affects what they buy; it affects what they do. It makes our job more difficult. [...] And I hope [...] that we can get that psychology turned around through persistence and disciplined policies*

Macro-Economic Realities in the 1970ies and 80ies

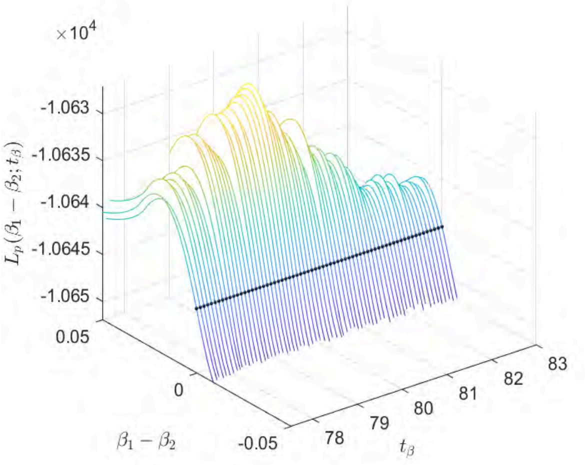


Rate of real interest (RRI), federal funds rate (FFP), consumer price index (CPI); source: Fed St.Louis

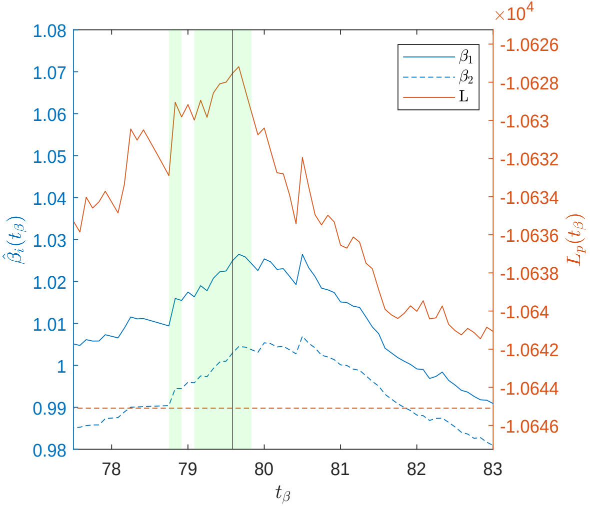
A Natural Experiment

- Conditions for a natural experiment
  - Volcker's announcement and implementation of his new monetary policy created a disruptive change in economic environment in 1979 and 1980.
  - The Zurcher data set sampling window is from 1974 to 1985
- Setup for a structural break model
  - Intuition: If there is a qualitative link between Zurcher's discounting and the real interest rates, we expect a reduction in discounting (due to a rise in real interest rates) shortly after Volcker's announcement.
  - Introduce an **unanticipated structural break** in Zurcher's discounting at an **unknown date**, t_β.
    - For $t < t_\beta$, Harold Zurcher discounts with β_1 from t up to ∞
    - For $t \geq t_\beta$, Harold Zurcher discounts with β_2 from t up to ∞
  - Test: $H_0 : \beta_1 = \beta_2$ against $H_0 : \beta_1 > \beta_2$ (or two-sided)

Tracing $\Delta\beta = \beta_1 - \beta_2$



The “Meeting” of Harold Zurcher and Paul Volcker



p	$\{1,2,3\},\{4,5,7\},\{6,8\}$
t_β	Sept 1979
	[Oct 1978, Dec 1978] $\bigcup$ [Feb 1979, Nov 1979]
β_1	1.0269
	$[1.0045, 1.0465]$
β_2	1.0047
	$[0.9852, 1.0197]$
LL	-10,627.1853
p-value ($H_0 : \beta_1 = \beta_2$)	$2.19 \cdot 10^{-9}$
p-value ($H_0 : \beta_1 = \beta_2 = 1$)	$6.55 \cdot 10^{-10}$

Conclusions and Outlook

- Methodological contribution
  - Method to *systematically* estimate models even in the presence of *non-identified* or *poorly identified* parameters
- Empirical contribution
  - Contrary to common belief, the discount factor in Rust (1987) is well identified and larger than 1
  - We find strong a strong indication for a qualitative link between the discount factor and the real interest rates as both the time of the structural break and the change of the discount rate agree
- Current work and outlook
  - Treat discounting as a *state* (exogenous, unobserved, serially correlated), and estimate the model using recursive likelihood function integration (RLI; Reich, 2018)
  - Apply this approach to model and data with *many* decision makers

References I

Aguirregabiria, V. and Mira, P. (2010). Dynamic Discrete Choice Structural Models: A Survey. *Journal of Econometrics*, 156(1):38–67.

Erdem, T. and Keane, M. (1996). Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science*, 15(1):1–20.

Heaton, J. (1995). An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications. *Econometrica: Journal of the Econometric Society*, 63(3):681.

Kocherlakota, N. R. (1990). On the 'discount' factor in growth economies. *Journal of Monetary Economics*, 25(1):43–47.

References II

Morton, T. E. and Wecker, W. E. (1977). Discounting, Ergodicity and Convergence for Markov Decision Processes. *Management Science*, 23(8):890–900.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica: Journal of the Econometric Society*, 55(5):999–1033.

Su, C.-L. and Judd, K. L. (2012). Constrained Optimization Approaches to Estimation of Structural Models. *Econometrica: Journal of the Econometric Society*, 80(5):2213–2230.

Rust (1987): Utility Function

- Agent's utility + shock for the single period payoff

$$u(x_t, d_t; \theta_{11}, \text{RC}) + \epsilon_t(d_t) = \begin{cases} -c(x_t, \theta_{11}) + \epsilon_t(0) & \text{if } d_t = 0 \\ -RC + \epsilon_t(1) & \text{if } d_t = 1 \end{cases}$$

  - $d_t = 0$: performing regular maintenance
  - $d_t = 1$: replacing the engine
- State variables
  - x_t mileage state
  - ϵ i.i.d. gumbel utility shock (only observed by agent)
- **Parameters**
  - θ_{11} regular maintenance cost parameter
  - RC replacement cost parameter

Value Function - Dynamic Programming

Objective The agent wants to maximize his expected discounted utility over an **infinite horizon**.

$$V_{\theta,\beta}(x_t, \epsilon_t) = \max_{D(x_t)\in\mathcal{D}} \mathbb{E}\left[\sum_{j=t}^{\infty} \beta^{j-t}\left(u(x_j, D(x_j); \theta_1, \mathrm{RC}) + \epsilon(D(x_j))\right) \middle| x_t\right]$$

where $\theta \equiv (\mathrm{RC}, \theta_1)$ and $D(\cdot)$ denotes the policy function.

⚠ For $\beta \to 1$, we have $V \to -\infty$

Bellman V is the unique solution to the Bellman equation

$$V_{\theta,\beta}(x, \epsilon) = \max_{d\in\{0,1\}} [u(x, d, \theta_1) + \epsilon(d) + \beta\mathbb{E}[V_{\theta,\beta}(x', \epsilon')|x, d]] =: T(V),$$

where x' and ϵ' denote the next period state variables.

Dynamic Programming with $\beta > 1$

Recall We can solve for the optimal value function by solving the fixed-point equation $V = T_{\theta,\beta}(V)$.

Note Even though, the value function $V \to \infty$, the **difference** between the value at different states might be finite

- We solve for the **relative value function** $h = V - V_1$ by

$$h = T_{\theta,\beta}(h) - T_{\theta,\beta}(h)_1$$

- For $\beta \in \mathbb{R}^+$, Morton & Wecker (1977) have derived a set of conditions under which the resulting policy is optimal

  Intuition: Relative values denote the *difference in total expected value* when starting from state x opposed to starting from state 1.

Structural Estimation

Data $\sim$15'000 observations of the state and control variables from *1974 to 1985*.

Objective Identify the most likely values for the parameters $\theta = (\theta_{11}, \mathrm{RC})$ and β given the observed data.

Approach Simultaneously solve the **likelihood** and **fixed-point** problem

$$\theta^*, \beta^* = \arg\max_{\theta,\beta} L(h, \theta, \beta; \{x_t, d_t\})$$

$$h = T_{\theta,\beta}(h) - T_{\theta,\beta}(h)_1|_{\theta=\theta^*,\beta=\beta^*}$$

$T(\cdot)$ denotes the Bellman operator, h the relative values

- Two popular solution methods are NFXP Rust (1987) and MPEC by Su and Judd (2012)

MPEC

- Su and Judd (2012) formulate the structural estimation as constrained optimization

$$\max_{(h,\theta,\beta)} L(\theta, \beta, h; \{x_t, d_t\}),$$
$$\text{s.t. } h = T_{\theta,\beta}(h) - T_{\theta,\beta}(h)_1,$$

with $\theta \in \mathbb{R}^2$, $\beta \in \mathbb{R}_+$, and $h \in \mathbb{R}^{90}$.

- β is "poorly identified" if the likelihood is (almost) flat in β. Numerically, its Hessian becomes (nearly) singular $\Rightarrow$ hard to estimate $\Rightarrow$ "calibrated" to some notion of the interest rate.

PMPEC

- We propose to formulate the structural estimation as parameterized MPEC:

$$L_p(\beta) = \max_{(h,\theta,\not\beta)} L(\theta, \beta, h; \{x_t, d_t\}),$$
$$\text{s.t. } h = T_{\theta,\beta}(h) - T_{\theta,\beta}(h)_1,$$

with $\theta \in \mathbb{R}^2$, $\beta \in \mathbb{R}_+$, and $h \in \mathbb{R}^{90}$.

- $L_p(\beta)$ denotes *augmented profile likelihood*

First-Order Necessary Optimality Conditions

- Its Lagrangian $\mathcal{L}$ is defined as

$$\mathcal{L}(\theta, h, \mu; \beta) = L(\theta, h; \beta) - \sum_i \mu_i \left(h - T_{\theta,\beta}(h) + T_{\theta,\beta}(h)_1\right)$$

- First-order necessary conditions form a system of equations **parameterized in $\boldsymbol{\beta}$**:

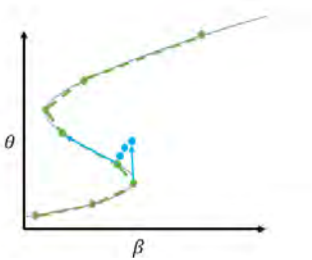
$$\nabla_{(\theta,h,\mu)} \mathcal{L}(\theta^*, h^*, \mu^*; \beta) = 0. \tag{1}$$

- We are interested in the solution manifold of the parametrized FOC (profile likelihood as implicit function)

$$c \equiv \{(\beta, \theta, h, \mu) : \nabla_{\theta,h,\mu} \mathcal{L}(\beta, \theta, h, \mu) = 0\}$$

Predictor–Corrector Homotopy Continuation

- Trace the one-dimensional solution manifold by *homotopy path continuation*.
- Predictor–Corrector in a nutshell



- Software: HOMPACK90 (Watson et al., 1997) with a self-written Matlab interface