Nonparametric Identification of Dynamic Models with Unobserved State Variables

Yingyao Hu and Matthew Shum

Johns Hopkins University

FEMES 2008

Singapore Management University

Introduction

- Consider identification of first-order Markov process $\{W_t, X_t^*\}$ for $t = 1, 2, ..., T$
- Only $\{W_t\}$ for $t = 1, 2, ..., T$ is observed
- Show:
  1. nonstationary: transition kernel $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$ identified from $f_{W_{t+1}, W_t, W_{t-1}, W_{t-2}, W_{t-3}}$ (5 obs.)
  2. stationary: transition kernel $f_{W_2, X_2^* | W_1, X_1^*}$ identified from $f_{W_{t+1}, W_t, W_{t-1}, W_{t-2}}$ (4 obs.)
- Identification of $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$ is crucial input for estimating dynamic models using "conditional-choice-probability (CCP)" approach of Hotz & Miller

Examples

- In most empirical applications, $W_t = (Y_t, M_t)$:
  - Y_t is "control variable": agent's action in period t
  - M_t is observed state variable
- X_t^* is persistent unobserved state variable
- **Example 1: generalized Rust (1987)**
  - Y_t: indicator for replacing bus engine
  - M_t: mileage of bus since last replacement
  - X_t^*: shocks to driver's ability, weather conditions, etc. (Rust assumed i.i.d over time)
- **Example 2: generalized Pakes (1986)**
  - Y_t: indicator for renewing patent
  - X_t^*: profitability from the patent (unobsd)
  - M_t: stock price, sales of patentholder

Roadmap

- Background
- Identification argument: discrete case (more details)
- Identification argument: continuous case (quickly)
- Simulation: 0-1 dichotomous case
- Example to illustrate assumptions: version of Rust (1987) bus engine replacement model
- Concluding remarks

Usefulness

We show identification of the joint density $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$; (also, unconditional $f_{W_t,X_t^*,W_{t-1},X_{t-1}^*}$ is identified).
In Markov dynamic choice models, this factorizes into economic components of interest:

$$\begin{aligned} f_{W_t,X_t^*|W_{t-1},X_{t-1}^*} &= f_{Y_t,M_t,X_t^*|Y_{t-1},M_{t-1},X_{t-1}^*} \\ &= \underbrace{f_{Y_t|M_t,X_t^*}}_{\text{CCP}} \cdot \underbrace{f_{M_t,X_t^*|Y_{t-1},M_{t-1},X_{t-1}^*}}_{\text{Markov state transitions}} \end{aligned}$$

From identified object, can recover: (i) conditional choice probability; (ii) Markov transitions for state variables.

Relation to literature

- Use as inputs into CCP-based approach to estimate dynamic discrete-choice model (Hotz-Miller, Aguirregabiria-Mira, Bajari-Benkard-Levin (2008), Pesendorfer-Schmidt-Dengler (2003), Pakes-Ostrovsky-Berry (2007)). Avoid numeric dyn. programming.
- First step in argument for nonparametric identification of DDC process with unobsd state variables (as in Magnac-Thesmar (2002), Bajari-Chernozhukov-Hong-Nekipelov (2005))
- Recent literature on estimating parametric DDC models with correlated USV
  - Bayesian: Imai, Jain, Ching (2006), Norets (2007)
  - Efficient simulation (particle filtering): Fernandez-Villaverde, Rubio-Ramirez (2006)

Relation to literature

- General criticism of CCP-based approaches: cannot accommodate unobservables which are persistent over time
- Recent literature on identification and estimation of DDC models with discrete and time-invariant X^* (unobserved heterogeneity)
  - Buchinsky-Hahn-Hotz (2004), Houde-Imai (2006), Kasahara-Shimotsu (2007)
  - Specifically: Kasahara-Shimotsu demonstrate identification of Markov process $W_t | W_{t-1}, X^*$
- Time-varying X_t^*:
  - Arcidiacono-Miller (2006): consider CCP estimation with discrete and time-varying X_t^*.
  - Cunha, Heckman, Schennach (2007): identify continuous X_t^* process in multivariate measurement error setting – W_t consists of noisy measurements of X_t^* and random noise

Relation to literature: nonclassical measurement errors

"Message": in X-section context, three "observations" (x, y, z) of latent x^* enough to identify (x, y, z, x^*)

- Hu (2008, JOE): X^*-discrete latent variable

$$f_{X,Y,Z} = \sum_{x^*} f_{X|X^*} f_{Y|X^*} f_{X^*,Z}$$

- Hu and Schennach (2008, ECMA): X^*:continuous latent variable

$$f_{X,Y,Z} = \int f_{X|X^*} f_{Y|X^*} f_{X^*,Z} dx^*$$

- Carroll, Chen and Hu (2008): $S-$sample indicator

$$f_{X,Y,Z,S} = \int f_{X|X^*,S} f_{Y|X^*,Z} f_{X^*,Z,S} dx^*$$

This paper

- X_t^* continuous
- X_t^* serially correlated: *unobserved state variable*
- Evolution of X_t^* can depend on W_{t-1}, X_{t-1}^*
- Focus on nonparametric identification of joint Markov process $W_t, X_t^* | W_{t-1}, X_{t-1}^*$

Basic setup: conditions for identification

- Consider dynamic processes $\{(W_T, X_T^*), ..., (W_t, X_t^*), ..., (W_1, X_1^*)\}_i$, i.i.d across agents $i \in \{1, 2, \ldots, n\}$.
- The researcher observes $\{W_{t+1}, W_t, W_{t-1}, W_{t-2}, W_{t-3}\}_i$ for many agents i (5 obs)
- Assumption: The dynamic process (W_t, X_t^*) satisfies

  (i) First-order Markov: $f_{W_t, X_t^* | W_{t-1}, \ldots, W_1, X_{t-1}^*, \ldots, X_1^*} = f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$

  (ii) Limited feedback: $f_{W_t | W_{t-1}, X_t^*, X_{t-1}^*} = f_{W_t | W_{t-1}, X_t^*}$. Picture

Comments on conditions

- Markov assumption standard in most applications of DDC models
- Limited feedback rules out direct effects from previous USV X^*_{t-1}, on current W_t. Implies that

$$\begin{aligned} f_{W_t|W_{t-1},X^*_t,X^*_{t-1}} &= f_{Y_t,M_t|Y_{t-1},M_{t-1},X^*_t,X^*_{t-1}} \\ &= f_{Y_t|M_t,Y_{t-1},M_{t-1},X^*_t,X^*_{t-1}} \cdot f_{M_t|Y_{t-1},M_{t-1},X^*_t,X^*_{t-1}} \\ &= \underbrace{f_{Y_t|M_t,Y_{t-1},M_{t-1},X^*_t}}_{\text{CCP}} \cdot \underbrace{f_{M_t|Y_{t-1},M_{t-1},X^*_t}}_{\text{mileage transition}} . \end{aligned}$$

- CCP usually simplifies further to $f_{Y_t|M_t,X^*_t}$.
- Simplification in mileage transition applies limited feedback condition. Satisfied by many empirical applications (in IO context: Crawford-Shum (2005), Das-Roberts-Tybout (2007), Xu (2008), Hendel-Nevo (2007))

Special case: Discrete X_t^*

- Main result for case of continuous X_t^*
- Build intuition by considering discrete case:

$$\forall t, \; X_t^* \in \mathcal{X}^* \equiv \{1, 2, \ldots, J\} .$$

- For convenience, assume W_t also discrete, with same support $\mathcal{W}_t = \mathcal{X}_t^*$.

Backbone of argument

For fixed (w_t, w_{t-1}), in matrix notation: here

$$L_{W_{t+1},w_t|w_{t-1},W_{t-2}} = L_{W_{t+1}|w_t,X_t^*} \cdot \underset{\displaystyle L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} \cdot L_{X_{t-1}^*|w_{t-1},W_{t-2}}}{\overset{\displaystyle L_{w_t,X_t^*|w_{t-1},W_{t-2}}}{\downarrow}}$$

Identify in several steps:

1&2: Get $f_{W_{t+1}|W_t,X_t^*}$

3: Get $f_{W_t,X_t^*|W_{t-1},W_{t-2}}$

4: Get $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$ (function of interest)

- BROWN: elements identified from data
- PURPLE: elements identified in proof

Step 1: identify $f_{W_{t+1}|W_t,X_t^*}$

- The key equation

$$
\begin{aligned}
& f_{W_{t+1},W_t,W_{t-1},W_{t-2}} \\
= & \int\int f_{W_{t+1},W_t,W_{t-1},W_{t-2},X_t^*,X_{t-1}^*} dx_t^* dx_{t-1}^* \\
= & \int\int f_{W_{t+1}|W_t,X_t^*} \cdot f_{W_t,X_t^*|W_{t-1},X_{t-1}^*} \cdot f_{W_{t-1},W_{t-2},X_{t-1}^*} dx_t^* dx_{t-1}^* \\
= & \int\int f_{W_{t+1}|W_t,X_t^*} \cdot f_{W_t|W_{t-1},X_t^*,X_{t-1}^*} \cdot f_{X_t^*,X_{t-1}^*,W_{t-1},W_{t-2}} dx_t^* dx_{t-1}^* \\
= & \int f_{W_{t+1}|W_t,X_t^*} f_{W_t|W_{t-1},X_t^*} \cdot f_{X_t^*,W_{t-1},W_{t-2}} dx_t^*
\end{aligned}
$$

- Discrete-case, matrix notation (for any fixed w_t, w_{t-1}) details:

$$
L_{W_{t+1},w_t|w_{t-1},W_{t-2}} = L_{W_{t+1}|w_t,X_t^*} D_{w_t|w_{t-1},X_t^*} L_{X_t^*|w_{t-1},W_{t-2}}
$$

Step 1 (cont'd)

- Important fact: for (w_t, w_{t-1}),

$$L_{W_{t+1},w_t|w_{t-1},W_{t-2}} = \underbrace{L_{W_{t+1}|w_t,X_t^*}}_{\text{no } w_{t-1}} \; \underbrace{D_{w_t|w_{t-1},X_t^*}}_{\text{only } J \text{ unkwns.}} \; \underbrace{L_{X_t^*|w_{t-1},W_{t-2}}}_{\text{no } w_t}$$

- for (w_t, w_{t-1}), $(\overline{w}_t, w_{t-1})$, $(\overline{w}_t, \overline{w}_{t-1})$ $(w_t, \overline{w}_{t-1})$,

$$\begin{aligned}
L_{W_{t+1},w_t|w_{t-1},W_{t-2}} &= L_{W_{t+1}|w_t,X_t^*} \; D_{w_t|w_{t-1},X_t^*} \; \underbrace{L_{X_t^*|w_{t-1},W_{t-2}}}_{\parallel} \\
L_{W_{t+1},\overline{w}_t|w_{t-1},W_{t-2}} &= \underbrace{L_{W_{t+1}|\overline{w}_t,X_t^*}}_{\parallel} \; D_{\overline{w}_t|w_{t-1},X_t^*} \; \overbrace{L_{X_t^*|w_{t-1},W_{t-2}}} \\
L_{W_{t+1},\overline{w}_t|\overline{w}_{t-1},W_{t-2}} &= \overbrace{L_{W_{t+1}|\overline{w}_t,X_t^*}} \; D_{\overline{w}_t|\overline{w}_{t-1},X_t^*} \; \underbrace{L_{X_t^*|\overline{w}_{t-1},W_{t-2}}}_{\parallel} \\
L_{W_{t+1},w_t|\overline{w}_{t-1},W_{t-2}} &= L_{W_{t+1}|w_t,X_t^*} \; D_{w_t|\overline{w}_{t-1},X_t^*} \; \overbrace{L_{X_t^*|\overline{w}_{t-1},W_{t-2}}}
\end{aligned}$$

Step 1: identify $f_{W_{t+1}|W_t,X_t^*}$

- Assume: LHS invertible, which is testable
- eliminate $L_{X_t^*|w_{t-1},W_{t-2}}$ using first two equations

$$\begin{aligned}\mathbf{A} &\equiv L_{W_{t+1},w_t|w_{t-1},W_{t-2}} L^{-1}_{W_{t+1},\overline{w}_t|w_{t-1},W_{t-2}} \\ &= L_{W_{t+1}|w_t,X_t^*} D_{w_t|w_{t-1},X_t^*} D^{-1}_{\overline{w}_t|w_{t-1},X_t^*} L^{-1}_{W_{t+1}|\overline{w}_t,X_t^*}\end{aligned}$$

- eliminate $L_{X_t^*|\overline{w}_{t-1},W_{t-2}}$ using last two equations

$$\begin{aligned}\mathbf{B} &\equiv L_{W_{t+1},w_t|\overline{w}_{t-1},W_{t-2}} L^{-1}_{W_{t+1},\overline{w}_t|\overline{w}_{t-1},W_{t-2}} \\ &= L_{W_{t+1}|w_t,X_t^*} D_{w_t|\overline{w}_{t-1},X_t^*} D^{-1}_{\overline{w}_t|\overline{w}_{t-1},X_t^*} L^{-1}_{W_{t+1}|\overline{w}_t,X_t^*}\end{aligned}$$

- eliminate $L^{-1}_{W_{t+1}|\overline{w}_t,X_t^*}$

$$\mathbf{AB}^{-1} = L_{W_{t+1}|w_t,X_t^*} D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*} L^{-1}_{W_{t+1}|w_t,X_t^*}$$

with diagonal matrix

$$D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*} = D_{w_t|w_{t-1},X_t^*} D^{-1}_{\overline{w}_t|w_{t-1},X_t^*} D_{\overline{w}_t|\overline{w}_{t-1},X_t^*} D^{-1}_{w_t|\overline{w}_{t-1},X_t^*}$$

Step 1: identify $f_{W_{t+1}|W_t,X_t^*}$

$$\mathbf{AB}^{-1} = L_{W_{t+1}|w_t,X_t^*} D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*} L^{-1}_{W_{t+1}|w_t,X_t^*}$$

represents eigenvalue-eigenvector decomposition of observed $\mathbf{AB}^{-1}$

- eigenvalues: diagonal entry in $D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*}$

$$\left(D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*}\right)_{j,j} = \frac{f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1},j)f_{W_t|W_{t-1},X_t^*}(\overline{w}_t|\overline{w}_{t-1},j)}{f_{W_t|W_{t-1},X_t^*}(\overline{w}_t|w_{t-1},j)f_{W_t|W_{t-1},X_t^*}(w_t|\overline{w}_{t-1},j)}$$

Assume: For uniqueness, $\left(D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*}\right)_{j,j}$ are finite, distinctive

- eigenvector: column in $L_{W_{t+1}|w_t,X_t^*}$, (normalized because sums to 1)

Hence, $L_{W_{t+1}|w_t,X_t^*}$ is identified (up to the value of x_t^*). Any permutation of eigenvectors yields same decomposition.

Step 1: identify $f_{W_{t+1}|W_t,X_t^*}$

To pin-down the value of x_t^*: need to "order" eigenvectors

- not necessary in the time-invariant case, $X_t^* = X_{t-1}^*$
- useful in time-varying case: show how agents change types w/ time.
- $f_{W_{t+1}|W_t,X_t^*}(\cdot|w_t, x_t^*)$ for any w_t is identified up to value of x_t^*
- To pin-down the value of x_t^* : Assume there is *known* functional

$$h(w_t, x_t^*) \equiv G\left[f_{W_{t+1}|W_t,X_t^*}(\cdot|w_t, \cdot)\right] \text{ is monotonic in } x_t^*.$$

  Then set $x_t^* = G\left[f_{W_{t+1}|W_t,X_t^*}(\cdot|w_t, \cdot)\right]$
- $G[f]$ may be mean, mode, median, other quantile of f.
- Note: in unobserved heterogeneity case ($X_t^* = X^*,\ \forall t$), it is enough to identify $f_{W_{t+1}|W_t,X_t^*}$.

Step 3: identify $f_{W_t, X_t^*, W_{t-1}, W_{t-2}}$

- Go back to main equation: for any (w_t, w_{t-1}) here

$$L_{W_{t+1}, w_t | w_{t-1}, W_{t-2}} = L_{W_{t+1} | w_t, X^*} \cdot L_{w_t, X_t^* | w_{t-1}, W_{t-2}}$$

- Identify $f_{W_t, X_t^* | W_{t-1}, W_{t-2}}$ through

$$L_{w_t, X_t^* | w_{t-1}, W_{t-2}} = L^{-1}_{W_{t+1} | w_t, X^*} \cdot L_{W_{t+1}, w_t | w_{t-1}, W_{t-2}}$$

- Also $f_{W_t, X_t^*, W_{t-1}, W_{t-2}} = f_{W_t, X_t^* | W_{t-1}, W_{t-2}} \cdot f_{W_{t-1}, W_{t-2}}$ known.

Step 4: identify $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$

- Markov property implies

$$f_{W_t,X_t^*|W_{t-1},W_{t-2}} = \sum_{X_{t-1}^* \in \mathcal{X}_{t-1}^*} f_{W_t,X_t^*|W_{t-1},X_{t-1}^*} \cdot f_{X_{t-1}^*|W_{t-1},W_{t-2}}$$

- Matrix notation (fixed w_t, w_{t-1}) here

$$L_{w_t,X_t^*|w_{t-1},W_{t-2}} = L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} L_{X_{t-1}^*|w_{t-1},W_{t-2}}$$

- Almost done, but what is $L_{X_{t-1}^*|w_{t-1},W_{t-2}}$? BUT: From

$$\begin{aligned} W_t, X_t^*, W_{t-1}, W_{t-2} \quad & \Rightarrow (\text{marginalize } W_{t-2}) \\ W_t, X_t^*, W_{t-1} \quad & = X_t^*|W_t, W_{t-1} \cdot W_t, W_{t-1} \end{aligned}$$

So we have $X_t^*|W_t, W_{t-1}$, but one-period off.

Step 4: identify $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$

- need 5 periods
- mimick above argument:
  From $f_{W_{t+1},W_t,W_{t-1},W_{t-2}}$ $\Rightarrow$ identify $f_{W_t,X_t^*,W_{t-1},W_{t-2}}$
  From $f_{W_t,W_{t-1},W_{t-2},W_{t-3}}$ $\Rightarrow$ identify $f_{W_{t-1},X_{t-1}^*,W_{t-2},W_{t-3}}$
- for any (w_t, w_{t-1})

$$
\begin{array}{ccccc}
L_{w_t,X_t^*|w_{t-1},W_{t-2}} & = & L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} & \cdot & L_{X_{t-1}^*|w_{t-1},W_{t-2}} \\
\Uparrow & & & & \Uparrow \\
f_{W_t,X_t^*,W_{t-1},W_{t-2}} & & & & f_{X_{t-1}^*|W_{t-1},W_{t-2}} \\
\Uparrow & & & & \Uparrow \\
f_{W_{t+1},W_t,W_{t-1},W_{t-2}} & & & & f_{W_{t-1},X_{t-1}^*,W_{t-2},W_{t-3}} \\
 & & & & \Uparrow \\
 & & & & f_{W_t,W_{t-1},W_{t-2},W_{t-3}}
\end{array}
$$

- Hence, $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$ is identified through

$$
L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} = L_{w_t,X_t^*|w_{t-1},W_{t-2}} L_{X_{t-1}^*|w_{t-1},W_{t-2}}^{-1} \blacksquare
$$

Stationary case

- stationarity: $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*} = f_{W_2, X_2^* | W_1, X_1^*}$
- only need 4 periods $\{W_{t+1}, W_t, W_{t-1}, W_{t-2}\}$
- stationarity helps identify $f_{X_{t-1}^* | W_{t-1}, W_{t-2}}$ without W_{t-3}

$$f_{W_t | W_{t-1}, W_{t-2}} = \int \underset{\substack{\| \\ f_{W_{t+1} | W_t, X_t^*} \\ \Uparrow \\ \text{step 1\&2}}}{f_{W_t | W_{t-1}, X_{t-1}^*}} f_{X_{t-1}^* | W_{t-1}, W_{t-2}} dx_{t-1}^*$$

- the rest is the same

Continuous case

- generalize the results in discrete case

discrete X_t^*	$\Rightarrow$	continuous X_t^*
matrix	$\Rightarrow$	linear operator here
invertible	$\Rightarrow$	one-to-one, "injective"
matrix diagonalization	$\Rightarrow$	spectral decomposition
eigenvector	$\Rightarrow$	eigenfunction

- $W_t = \mathcal{W}_t \subseteq \mathbb{R}^d$, $X_t^* \in \mathcal{X}_t^* \subseteq \mathbb{R}$, for all t
- Assume known fxn to reduce W_t to same dimensionality as X_t^*:

$$V_t = g(W_t), \text{ where } g : \mathcal{W}_t \to \mathcal{X}_t^*$$

For convenience: avoid complicated argument involving adjoint operators (no extra insights)

- Example: Step 1

Summary of assumptions

1. (i) First-order Markov $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*,\Omega_{<t-1}} = f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$;
   (ii) Limited feedback $f_{W_t|W_{t-1},X_t^*,X_{t-1}^*} = f_{W_t|W_{t-1},X_t^*}$
2. (Invertibility) for any w_t, w_{t-1}, (i) $L_{V_{t+1}|w_t,X_t^*}$ is one-to-one ;
   (ii) $L_{V_{t+1},w_t|w_{t-1},V_{t-2}}$ is one-to-one
3. (finite, distinctive eigenvalues) for any w_t, (i)

$$0 < L(w_t, w_{t-1}) \leq f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1}, x_t^*) \leq U(w_t, w_{t-1}) < \infty$$

   (ii) for any x_t^* and w_t, there exists w_{t-1} such that

$$\frac{\partial^3}{\partial w_t \partial w_{t-1} \partial x_t^*} \ln f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1}, x_t^*) \neq 0.$$

4. (normalize value of x_t^*) for any $w_t \in \mathcal{W}_t$,
   $x_t^* = G\left\{f_{V_{t+1}|W_t,X_t^*}(\cdot|w_t, x_t^*)\right\}$
5. For any $w_{t-1} \in \mathcal{W}_{t-1}$, $L_{X_{t-1}^*|w_{t-1},V_{t-2}}$ is one-to-one. NB

Main result

- **Theorem 1**: Under assumptions above, the density $f_{W_{t+1},W_t,W_{t-1},W_{t-2},W_{t-3}}$ uniquely determines $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$
- **Theorem 2**: With stationarity, the density $f_{W_{t+1},W_t,W_{t-1},W_{t-2}}$ uniquely determines $f_{W_2,X_2^*|W_1,X_1^*}$
- We can use existing argument from Magnac-Thesmar, Bajari-Chernozhukov-Hong-Nekipelov to argue identification of utility functions, once $W_t, X_t^*|W_{t-1}, X_{t-1}^*$ known here

Simulation

- exactly follow the identification procedure of nonstationary case
- $\{W_t, X_t^*\}$ is generated as follows: $u_1, u_2 \sim uniform\,(0,1)$

$$W_t = \begin{cases} I\,(u_1 > 0.95) & \text{if } (X_t^*, W_{t-1}) = (0,0) \\ I\,(u_1 > 0.60) & \text{if } (X_t^*, W_{t-1}) = (1,0) \\ I\,(u_1 > 0.05) & \text{if } (X_t^*, W_{t-1}) = (0,1) \\ I\,(u_1 > 0.50) & \text{if } (X_t^*, W_{t-1}) = (1,1) \end{cases},$$

$$X_t^* = \begin{cases} I\,(u_2 > 0.25) & \text{if } \left(X_{t-1}^*, W_{t-1}\right) = (0,0) \\ I\,(u_2 > 0.75) & \text{if } \left(X_{t-1}^*, W_{t-1}\right) = (1,0) \\ I\,(u_2 > 0.60) & \text{if } \left(X_{t-1}^*, W_{t-1}\right) = (0,1) \\ I\,(u_2 > 0.05) & \text{if } \left(X_{t-1}^*, W_{t-1}\right) = (1,1) \end{cases}.$$

- two estimators: using $\{W_t\}$ and using $\{W_t, X_t^*\}$
- n=50000, reps=200: $\Longrightarrow$ mean (std.err)

Simulation

$\widehat{f}(W_t, X_t^* \vert W_{t-1}, X_{t-1}^*)$	using $\{W_t\}$	using $\{W_t, X_t^*\}$	mean Differ.
(0 , 0 | 0 , 0)	0.0454 (0.0754)	0.0475 (0.0019)	-0.0021
(0 , 0 | 0 , 1)	0.4768 (0.0499)	0.4752 (0.0032)	0.0016
(0 , 0 | 1 , 0)	0.1357 (0.1354)	0.1491 (0.0075)	-0.0134
(0 , 0 | 1 , 1)	0.0030 (0.0092)	0.0011 (0.0008)	0.0019
(0 , 1 | 0 , 0)	0.5543 (0.0501)	0.5703 (0.0046)	-0.0161
(0 , 1 | 0 , 1)	0.2985 (0.0453)	0.3000 (0.0030)	-0.0015
(0 , 1 | 1 , 0)	0.3008 (0.1341)	0.3002 (0.0100)	0.0006
(0 , 1 | 1 , 1)	0.7317 (0.0136)	0.7465 (0.0047)	-0.0148
(1 , 0 | 0 , 0)	0.0021 (0.0047)	0.0025 (0.0004)	-0.0004
(1 , 0 | 0 , 1)	0.0245 (0.0176)	0.0250 (0.0011)	-0.0005
(1 , 0 | 1 , 0)	0.4363 (0.0886)	0.4504 (0.0103)	-0.0142
(1 , 0 | 1 , 1)	0.0083 (0.0210)	0.0033 (0.0024)	0.0050
(1 , 1 | 0 , 0)	0.3716 (0.0212)	0.3797 (0.0045)	-0.0081
(1 , 1 | 0 , 1)	0.1992 (0.0189)	0.1998 (0.0028)	-0.0006
(1 , 1 | 1 , 0)	0.1007 (0.0453)	0.1002 (0.0068)	0.0004
(1 , 1 | 1 , 1)	0.2441 (0.0143)	0.2491 (0.0040)	-0.0049

Discuss assumptions: example from Rust (1987)

Consider particular version of Rust (1987): $W_t = (Y_t, M_t)$:

- $Y_t \in \{0, 1\}$ (don't replace, replace)
- M_t is mileage
- X_t^* is trunc. normal process w/ bounded support $[L, U]$:

$$X_t^* = \begin{cases} 0.5X_{t-1}^* + 0.3\psi(M_{t-1}) + 0.2\nu_t & \text{if } Y_{t-1} = 0 \\ 0.8X_{t-1}^* + 0.2\nu_t & \text{if } Y_{t-1} = 1 \end{cases}$$

  - ν_t are i.i.d. truncated normal on $[L, U]$.
  - $\psi(M_{t-1}) = L + (U - L)\frac{\exp(M_{t-1})-1}{\exp(M_{t-1})+1}$,
- Dimension-redxn: $V_t = g(W_t) = M_t$ (continuous element of W_t)

Two different specifications:

Specification A	Specification B
$u_t = \begin{cases} -c(M_t) + X_t^* + \epsilon_{0t}, & Y_t = 0 \\ -RC + \epsilon_{1t}, & Y_t = 1. \end{cases}$ $c(\cdot)$ bounded away from $0, +\infty$	$u_t = \begin{cases} -c(M_t) + \epsilon_{0t} \\ -RC + \epsilon_{1t} \end{cases}$ ...
$M_{t+1} = \begin{cases} M_t + \eta_{t+1}, & Y_t = 0 \\ \eta_{t+1}, & Y_t = 1 \end{cases}$ η_t are $N(0,1)$, trunc. to $[0,1]$, i.i.d.	$M_{t+1} = \begin{cases} M_t + \eta_{t+1} \cdot \exp(X_{t+1}^*) \\ \eta_{t+1} \cdot \exp(X_{t+1}^*). \end{cases}$ ...

- Specifications differ in where X_t^* enters.
- Discuss each assumption in turn
- Assumption 1 (Markov, LF) satisfied

Assumption 2

- $L_{M_{t+1},w_t|w_{t-1},M_{t-2}}$ is one-to-one:
  Consider w_t where $y_t = 1$.
  - **A:** M_{t+1} is trunc. $N(0,1)$, regardless of (w_{t-1}, M_{t-2}). FAILS
  - **B:** M_{t+1} depends on X^*_{t+1}, which is correlated with M_{t-2}. OK
- $L_{M_{t+1}|w_t,X^*_t}$ is one-to-one:
  Again, consider w_t where $y_t = 1$.
  - **A:** $M_{t+1}|w_t, X^*_t$ is trunc. $N(0,1)$. FAILS
  - **B:** $M_{t+1}|w_t, X^*_t$ depends on X^*_t. OK
- Note: One-to-one rules out models where W_t only has discrete components, but X^*_t is continuous.

Assumption 3: Finite, distinct eigenvalues

1. Sufficient cdtn for *finite eigenvalues*: for all (w_t, w_{t-1}), there exist functions $L(w_t, w_{t-1})$, $U(w_t, w_{t-1})$ st for all x_t^*:

$$0 < L(w_t, w_{t-1}) \leq f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1}, x_t^*) \leq U(w_t, w_{t-1}) < \infty.$$

- $f_{W_t|W_{t-1},X_t^*} = f_{Y_t|M_t,X_t^*} \cdot f_{M_t|X_t^*,Y_{t-1},M_{t-1}}$.
  Are all terms bounded away from 0, $+\infty$?
  - $f_{M_t|X_t^*,Y_{t-1},M_{t-1}}$ is truncated $N(0,1)$. OK
  - Per-period utilities bounded (except ϵ's), so CCP's also bounded away from 0
- Boundedness assumptions on M_t, period utility functions without much loss of generality. (Usually good for computing models)

Assumption 3: cont'd

2. Sufficient cdtn for *distinct eigenvalues*: for any $x_t^* \in \mathcal{X}_t^*$ & $w_t \in \mathcal{W}_t$, there exists $w_{t-1} \in \mathcal{W}_{t-1}$ st

$$\frac{\partial^3}{\partial m_t \partial m_{t-1} \partial x_t^*} \ln f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1},x_t^*) \neq 0.$$

Spec. B: pick w_{t-1} st $y_{t-1} = 0$.

$$m_t|m_{t-1}, y_{t-1}, X_t^* \sim \frac{1}{\exp(X_t^*)} \cdot \tilde{\phi}\left(\frac{m_t - m_{t-1}}{\exp(X_t^*)}\right)$$

where $\tilde{\phi}(\cdot)$ is N(0,1) density truncated to [0,1].
Clearly, $\frac{\partial^3}{\partial m_t \partial m_{t-1} \partial x_t^*} \ln f_{M_t|X_t^*,Y_{t-1},M_{t-1}}(m_t|m_{t-1},y_{t-1},X_t^*) \neq 0$, implying sufficient cdtn.
Spec. A: $m_t|m_{t-1}, y_{t-1}, X_t^*$ is never function of X_t^*. Sufficient cdtn cannot hold.

Assumption 4

Appropriate normalization to pin down unobserved X_t^*

- Median of $f_{M_{t+1}|M_t,Y_t,X_t^*}(\cdot|m_t, y_t, z)$ is

$$h(w_t, z) = \begin{cases} m_t + C_{med} \cdot \exp(0.3\psi(m_t)) \cdot \exp(0.5z) & \text{if } y_t = 0 \\ C_{med} \cdot \exp(0.8z) & \text{if } y_t = 1, \end{cases}$$

where C_{med} denotes med $\left[\eta_{t+1} \cdot \exp(0.2\nu_{t+1})\right]$ (fixed).

- $h(w_t, z)$ is monotonic in z
- So pin down $x_t^* = med\left[f_{M_{t+1}|M_t,Y_t,X_t^*}(\cdot|m_t, y_t, x_t^*)\right]$

Assumption 5

$L_{X^*_{t-1}|w_{t-1},M_{t-2}}$ is one-to-one (from M_{t-2} to X^*_{t-1}).

- From inspection of transition density for the latent process X^*_t: X^*_{t-1} depends on M_{t-2} if $Y_{t-2} = 0$, but not if $Y_{t-2} = 1$.
- Conditional distribution of $X^*_{t-1}|w_{t-1}, M_{t-2}$ includes observations with both $Y_{t-2} = 1, 0$.
- So long as $P(Y_{t-2} = 0|w_{t-1}, M_{t-2}) > 0$, then one-to-one assumption should hold.

Concluding remarks

- Identification of Markov process $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$, where X_t^* is unobserved state variable
  1. nonstationary: transition kernel $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$ identified from $f_{W_{t+1}, W_t, W_{t-1}, W_{t-2}, W_{t-3}}$ (5 obs.)
  2. stationary: transition kernel $f_{W_2, X_2^* | W_1, X_1^*}$ identified from $f_{W_{t+1}, W_t, W_{t-1}, W_{t-2}}$ (4 obs.)
- Ongoing work: apply these results to estimate DDC models with unobserved state variables.
  Start with discrete X_t^* case: identification proofed mimicked for estimation. Continuous case harder (invert linear operators) here
- Extension: allow latent process X_t^* to be multivariate. Useful for dynamic games applications (X_t^* includes USV's for each player).

Estimation

- Recall that

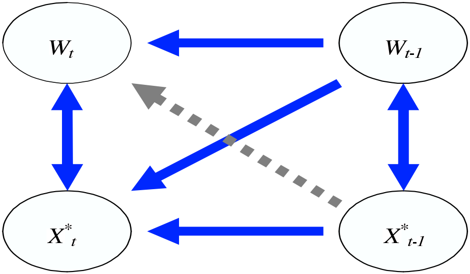
$$f_{W_t,X_t^*|W_{t-1},X_{t-1}^*} = f_{W_t|W_{t-1},X_t^*} \cdot f_{X_t^*|X_{t-1}^*,W_{t-1}}.$$

- We have shown $f_{W_{t+1}|W_t,X_t^*}$, $f_{W_t|W_{t-1},X_t^*}$, and $f_{X_t^*|X_{t-1}^*,W_{t-1}}$ are identified from
$f_{W_{t+1},W_t,W_{t-1},W_{t-2},W_{t-3}} =$
$\int f_{W_{t+1}|W_t,X_t^*} f_{W_t|W_{t-1},X_t^*} \left(\int f_{X_t^*|X_{t-1}^*,W_{t-1}} f_{X_{t-1}^*,W_{t-1},W_{t-2},W_{t-3}} dx_{t-1}^* \right) dx_t^*$
- Leads to a semi-nonparametric MLE based on this density

Return

Flowchart

Return



Identifying utility functions (sketch)

Assumptions:

1. Action set: $\mathcal{Y} = \{0, 1, \ldots, K\}$.
2. State variables are $S \equiv (M, X^*)$.
3. Per-period utility from choosing $y \in \mathcal{Y}$:

$$u_y(S_t) + \epsilon_{y,t}, \ \forall y \in \mathcal{Y}, \ \epsilon \sim F(\epsilon), \ i.i.d.$$

4. From data, the CCP's $P_y(S) \equiv \text{Prob}(Y = 1|S)$ and state transitions $p(S'|Y, S)$ are identified. (Main Theorem)
5. $u_0(S) = 0$, for all S
6. Discount factor β is known.

Goal: From $W', X^{*\prime}|W, X^*$, identify $u_y(\cdot), \ y = 1, \ldots, K$

Identifying utility functions

- From HM, MT: $\exists$ *known* one-to-one mapping $q(S) : \mathbb{R}^K \rightarrow \mathbb{R}^K$, which maps $(p_1(S), \ldots, p_K(S))$ to $(\Delta_1(S), \ldots, \Delta_K(S))$, where

  $$\Delta_y(S) \equiv V_y(S) - V_0(S) \text{ diff. in choice-specific value functions.}$$

- "Bellman" equation for zero choice:

  $$V_0(S) = \beta E_{S'|S,Y}\left[G(\Delta_1(S'), \ldots, \Delta_K(S')) + V_0(S')\right].$$

  Hence, can recover $V_0(\cdot)$ function. G is "social-surplus" function (known).

- Hence, utilities identified from

  $$u_y(S) = V_y(S) - \beta E_{S'|S,Y}\left[G(\Delta_1(S'), \ldots, \Delta_K(S')) + V_0(S')\right], \forall y \in \mathcal{Y},$$

Logit case | Return

Linear operators

- for example, for given w_t, w_{t-1}

$$\left(L_{W_{t+1}|w_t,X_t^*}h\right)(x) = \int f_{W_{t+1}|W_t,X_t^*}(x|w_t,x_t^*)h(x_t^*)dx_t^*$$

$$\left(L_{W_{t+1},w_t|w_{t-1},X_{t-2}}h\right)(x) = \int f_{W_{t+1},W_t|W_{t-1},X_{t-2}}\left(x,w_t|w_{t-1},z\right)h(z)dz.$$

- Matrix is linear operator in finite-dimensional space

Return

Continuous case: Step 1

Return

- The key equation is

$$f_{V_{t+1},W_t|W_{t-1},V_{t-2}} = \int f_{V_{t+1}|W_t,X_t^*} f_{W_t|W_{t-1},X_t^*} f_{X_t^*|W_{t-1},V_{t-2}} dx_t^*.$$

- decomposition of an observed operator

$$L_{V_{t+1},w_t|w_{t-1},V_{t-2}} L^{-1}_{V_{t+1},\overline{w}_t|w_{t-1},V_{t-2}} \left(L_{V_{t+1},w_t|\overline{w}_{t-1},V_{t-2}} L^{-1}_{V_{t+1},\overline{w}_t|\overline{w}_{t-1},V_{t-2}} \right)^{-1}$$
$$= L_{V_{t+1}|w_t,X_t^*} D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*} L^{-1}_{V_{t+1}|w_t,X_t^*}$$

where a diagonal operator $D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*}$:

$$\left(D_{w_t,\overline{w}_t,w_{t-1},\overline{w}_{t-1},X_t^*} g\right)(x_t^*) = k\left(w_t, \overline{w}_t, w_{t-1}, \overline{w}_{t-1}, x_t^*\right) g(x_t^*).$$

- eigenvalue for index x_t^*

$$k\left(\dots, x_t^*\right) = \frac{f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1}, x_t^*) f_{W_t|W_{t-1},X_t^*}(\overline{w}_t|\overline{w}_{t-1}, x_t^*)}{f_{W_t|W_{t-1},X_t^*}(\overline{w}_t|w_{t-1}, x_t^*) f_{W_t|W_{t-1},X_t^*}(w_t|\overline{w}_{t-1}, x_t^*)}.$$

- eigenfunction for index x_t^* : $f_{V_{t+1}|W_t,X_t^*}(\cdot|w_t, x_t^*)$

Further details on assumption 5

$$L^{-1}_{X_{t+1}|w_t,X_t^*} L_{V_{t+1},w_t|w_{t-1},V_{t-2}} = L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} \cdot L_{X_{t-1}^*|w_{t-1},V_{t-2}}$$

$$\Uparrow$$

$$D_{w_t|w_{t-1},X_t^*} L_{X_t^*|w_{t-1},V_{t-2}}$$

$$\Uparrow$$

$$L_{w_t,X_t^*|w_{t-1},V_{t-2}}$$

- By Assumption 2, both LHS operators are one-to-one
- By Assumption 5, second operator on RHS is one-to-one
- Hence, we can conclude

$$L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} = L^{-1}_{V_{t+1}|w_t,X_t^*} L_{V_{t+1},w_t|w_{t-1},V_{t-2}} L^{-1}_{X_{t-1}^*|w_{t-1},V_{t-2}}$$

Return

Matrix definitions

- $L_{w_t,X_t^*,w_{t-1},W_{t-2}} = \left[f_{W_t,X_t^*,W_{t-1},W_{t-2}}(w_t,i|w_{t-1},j)\right]_{i,j}$ Return 2
-

$$\begin{aligned} L_{w_t,X_t^*|w_{t-1},W_{t-2}} &= \left[f_{W_t,X_t^*|W_{t-1},W_{t-2}}(w_t,i|w_{t-1},j)\right]_{i,j} \\ L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} &= \left[f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}(w_t,i|w_{t-1},j)\right]_{i,j} \\ L_{X_{t-1}^*|w_{t-1},W_{t-2}} &= \left[f_{X_{t-1}^*|W_{t-1},W_{t-2}}(i|w_{t-1},j)\right]_{i,j} \end{aligned}$$

Return 3

For w_t, set $\mathcal{B}(w_t)$ contains points $(\overline{w}_t, w_{t-1}, \overline{w}_{t-1})$ satisfying:

1. $\overline{w}_t \in \mathcal{W}_t$; $w_{t-1}, \overline{w}_{t-1} \in \mathcal{A}(w_t) \cap \mathcal{A}(\overline{w}_t)$; $\overline{w}_t \neq w_t$; and $\overline{w}_{t-1} \neq w_{t-1}$; Implies that $L_{X_{t+1},\overline{w}_t|w_{t-1},Z_{t-2}}$, $L_{X_{t+1},w_t|\overline{w}_{t-1},Z_{t-2}}$, $L_{X_{t+1},\overline{w}_t|\overline{w}_{t-1},Z_{t-2}}$ are 1-to-1
2. $k(w_t, \overline{w}_t, w_{t-1}, \overline{w}_{t-1}, x_t^*) < \infty$ for all $x_t^* \in \mathcal{X}_t^*$.
   So AB^{-1} is bounded operator.
   Sufficient condition for $k(w_t, \overline{w}_t, w_{t-1}, \overline{w}_{t-1}, x_t^*) < \infty$ for all $x_t^* \in \mathcal{X}_t^*$: for all $(w_t, w_{t-1}) \in \mathcal{W}_t \times \mathcal{W}_{t-1}$, $\exists$ $L(w_t, w_{t-1})$ and $U(w_t, w_{t-1})$ st for all x_t^*

$$0 < L(w_t, w_{t-1}) \leq f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1}, x_t^*) \leq U(w_t, w_{t-1}) < \infty.$$

Return

Logit case

$$G(\Delta_1(S), \ldots, \Delta_K(S)) = \log\left[1 + \sum_{y=1}^{K} \exp(\Delta_y(S))\right]$$

$$q_y(S) = \Delta_y(S) = \log(p_y(S)) - \log(p_0(S)), \ \forall y = 1, \ldots K,$$

where $p_0(S) \equiv 1 - \sum_{y=1}^{K} p_y(S)$.

Return

Matrix notation

- Define the J-by-J matrices (fix w_t and w_{t-1})

$$
\begin{aligned}
L_{W_{t+1},w_t|w_{t-1},W_{t-2}} &= \left[f_{W_{t+1},W_t|W_{t-1},W_{t-2}}(i, w_t|w_{t-1}, j)\right]_{i,j} \\
L_{W_{t+1}|w_t,X_t^*} &= \left[f_{W_{t+1}|W_t,X_t^*}(i|w_t, j)\right]_{i,j} \\
L_{X_t^*|w_{t-1},W_{t-2}} &= \left[f_{X_t^*|W_{t-1},W_{t-2}}\left(i|w_{t-1}, j\right)\right]_{i,j}
\end{aligned}
$$

$$
D_{w_t|w_{t-1},X_t^*} = \left[\begin{array}{ccc} f_{W_t|W_{t-1},X_t^*}\left(w_t|w_{t-1},1\right) & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & f_{W_t|W_{t-1},X_t^*}\left(w_t|w_{t-1},J\right) \end{array}\right]
$$

Return1

Return2