Nonparametric Identification of Dynamic Models with Unobserved State Variables

Yingyao Hu and Matthew Shum

Johns Hopkins University & Caltech

- **Example 1: Dynamic Investment Model**
  - Y_t: firm investment
  - M_t: capital stock
  - X_t^*: firm-level productivity
- **Example 2: Dynamic learning**
  - Y_t: which brand is consumed
  - M_t: # advertisements seen
  - X_t^*: current beliefs (posterior mean) about each brand
- Examples of **Markov dynamic choice models** with *serially correlated unobserved state variables*

Introduction

Data problem:

- Consider identification of first-order Markov process $\{W_t, X_t^*\}_{t=1}^T$
- Only $\{W_t\}$ for $t = 1, 2, ..., T$ is observed
  - In most empirical dynamic models, $W_t = (Y_t, M_t)$:
    - Y_t is choice variable: agent's action in period t
    - M_t is observed state variable
  - X_t^* is persistent (serially-correlated) unobserved state variable
- In these models, structural components fully summarized in Markov law of motion $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$.
  $\Rightarrow$ study nonparametric identification of this.

Two main results:

1. **Nonstationary case**: for each period t, the law of motion $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$ identified from *five* observations $W_{t+1}, W_t, W_{t-1}, W_{t-2}, W_{t-3}$
2. **Stationary case**: Markov law of motion $f_{W_2,X_2^*|W_1,X_1^*}$ identified from *four* observations $W_{t+1}, W_t, W_{t-1}, W_{t-2}$

Usefulness

Once we identify $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$, it factorizes into structural components of interest:

$$f_{W_t,X_t^*|W_{t-1},X_{t-1}^*} = f_{Y_t,M_t,X_t^*|Y_{t-1},M_{t-1},X_{t-1}^*} = \underbrace{f_{Y_t|M_t,X_t^*}}_{\text{CCP}} \cdot \underbrace{f_{M_t,X_t^*|Y_{t-1},M_{t-1},X_{t-1}^*}}_{\text{Markov state laws of motion}}$$

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

Once these are known, can estimate "structural" parameters of model (eg. utility parameters) using "conditional-choice-probability (CCP)" pioneered by Hotz & Miller

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

Relation to literature

- 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), Hong-Shum (2007)). Virtue: avoid numeric dyn. programming.
- Empirical applications include: Ryan (2006), Collard-Wexler (2006)
- Nonparametric identification of DDC models (as in Magnac-Thesmar (2002), Bajari-Chernozhukov-Hong-Nekipelov (2005))
- General criticism of CCP-based approaches: cannot accommodate unobservables which are persistent over time $\Longrightarrow$

Recent literature

- Dynamic models with time-invariant X^* (unobsd het)
  - Buchinsky-Hahn-Hotz (2004), Houde-Imai (2006)
  - Kasahara-Shimotsu (2007): identify Markov process $W_t|W_{t-1}, X^*$
- Time-varying X_t^*:
  - Arcidiacono-Miller (2006): CCP estimation; discrete, time-varying X_t^*.
  - Henry, Kitamura, Salanie (2008): identification in dynamic, discrete "hidden-Markov state" models
  - Cunha, Heckman, Schennach (2007): multivariate measurement error setting – unobserved *process* $\{X_t^*\}_{t=1}^T$ noisily measured by $\{W_{1t}\}_{t=1}^T$, $\{W_{2t}\}_{t=1}^T$, $\{W_{3t}\}_{t=1}^T$, *cond. indep.*.
- Estimating parametric DDC models w/ correlated USV
  - Bayesian: Imai, Jain, Ching (2006), Norets (2007), Gallant-Hong-Khwaja (2008)
  - Efficient simulation: Fernandez-Villaverde, Rubio-Ramirez (2006), Blevins (2008)

Our contribution

- 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 $W_t, X_t^* | W_{t-1}, X_{t-1}^*$
- Novel identification approach: use recent findings from nonclassical measurement error econometrics: Hu (2008), Hu-Schennach (2007), Carroll, Chen, and Hu (2008)

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 (*this paper*)

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

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: $f_{W_t, X_t^* | W_{t-1}, ..., W_1, X_{t-1}^*, ..., X_1^*} = f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$

  (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
- LF rules out direct effects from previous X^*_{t-1} to W_t:

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

- LF restricts M_t law of motion.
  Satisfied by many empirical applications (in IO context: Crawford-Shum (2005), Das-Roberts-Tybout (2007), Xu (2008), Hendel-Nevo (2007)) Details
- To relax LF: (i) impose additional restrictions on CCP; (ii) identify higher-order Markov models (in progress)

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 $\mathcal{W}_t = \mathcal{X}_t^*$.
- In what follows:
  - "L" denotes J-square matrix
  - "D" denotes J-diagonal matrix.

Backbone of argument

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

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

- **Lemma 2:** Markov law of motion $L_{w_t, X_t^* | w_{t-1}, X_{t-1}^*}$

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

  Hence, all we must identify are $L_{W_{t+1}|w_t, X_t^*}$ and $L_{W_t | w_{t-1}, X_{t-1}^*}$.
- **Lemma 3:** From $f_{W_{t+1}, w_t | w_{t-1}, W_{t-2}}$, identify $L_{W_{t+1}|w_t, X_t^*}$.
- Stationary case: $L_{W_{t+1}|w_t, X_t^*} = L_{W_t | w_{t-1}, X_{t-1}^*}$, so Lemma 3 implies identification (4 obs)
- Non-stationary case: apply Lemma 3 in turn to $f_{W_{t+1}, w_t | w_{t-1}, W_{t-2}}$ and $f_{W_t, w_{t-1} | w_{t-2}, W_{t-3}}$ (5 obs)

Lemma 2: representation of $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$

- Main equation: for any (w_t, w_{t-1}) here here

$$
\begin{aligned}
L_{W_{t+1},w_t|w_{t-1},W_{t-2}} &= L_{W_{t+1}|w_t,X^*} L_{w_t,X_t^*|w_{t-1},W_{t-2}} \\
&= L_{W_{t+1}|w_t,X^*} L_{w_t,X_t^*|w_{t-1},X_{t-1}^*} L_{X_{t-1}^*|w_{t-1},W_{t-2}}
\end{aligned}
$$

- Similarly: $L_{W_t|w_{t-1},W_{t-2}} = L_{W_t|w_{t-1},X_{t-1}^*} L_{X_{t-1}^*|w_{t-1},W_{t-2}}$
- Manipulating above two equations: $L_{w_t,X_t^*|w_{t-1},X_{t-1}^*}$

$$
\begin{aligned}
&= L^{-1}_{W_{t+1}|w_t,X^*} L_{W_{t+1},w_t|w_{t-1},W_{t-2}} L^{-1}_{X_{t-1}^*|w_{t-1},W_{t-2}} \\
&= L^{-1}_{W_{t+1}|w_t,X_t^*} L_{W_{t+1},w_t|w_{t-1},W_{t-2}} L^{-1}_{W_t|w_{t-1},W_{t-2}} L_{W_t|w_{t-1},X_{t-1}^*}
\end{aligned}
$$

- Identification of $L_{w_t,X_t^*|w_{t-1},X_{t-1}^*}$ boils down to that of $L_{W_{t+1}|w_t,X_t^*}$ for t & $t-1$ (Lemma 3)

Lemma 3: proof

- Similar to Carroll, Chen, and Hu (2008)
- The key equation: $f_{W_{t+1},W_t,W_{t-1},W_{t-2}}$

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

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

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

Lemma 3: Proof (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}), $(\overline{w}_t, w_{t-1})$, $(\overline{w}_t, \overline{w}_{t-1})$ $(w_t, \overline{w}_{t-1})$,

$$L_{W_{t+1},w_t|w_{t-1},W_{t-2}} = L_{W_{t+1}|w_t,X_t^*} \quad D_{w_t|w_{t-1},X_t^*} \quad \underbrace{L_{X_t^*|w_{t-1},W_{t-2}}}$$
$$\|$$
$$L_{W_{t+1},\overline{w}_t|w_{t-1},W_{t-2}} = \underbrace{L_{W_{t+1}|\overline{w}_t,X_t^*}} \quad D_{\overline{w}_t|w_{t-1},X_t^*} \quad \overbrace{L_{X_t^*|w_{t-1},W_{t-2}}}$$
$$\|$$
$$L_{W_{t+1},\overline{w}_t|\overline{w}_{t-1},W_{t-2}} = \overbrace{L_{W_{t+1}|\overline{w}_t,X_t^*}} \quad D_{\overline{w}_t|\overline{w}_{t-1},X_t^*} \quad \underbrace{L_{X_t^*|\overline{w}_{t-1},W_{t-2}}}$$
$$\|$$
$$L_{W_{t+1},w_t|\overline{w}_{t-1},W_{t-2}} = L_{W_{t+1}|w_t,X_t^*} \quad D_{w_t|\overline{w}_{t-1},X_t^*} \quad \overbrace{L_{X_t^*|\overline{w}_{t-1},W_{t-2}}}$$

Lemma 3: Proof (cont'd)

- 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

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

Lemma 3: Proof (cont'd)

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

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

- 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.

Lemma 3: Proof (cont'd)

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 $f_{W_{t+1}|W_t,X_t^*}$.

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
- Example: Step 1

Assumptions

1. (i) First-order Markov; (ii) Limited feedback
2. (Invertibility) There exists variable(s) $V \subseteq W$ st, for any w_t, w_{t-1}, the following are one-to-one: (i) $L_{V_{t-2}, w_t | w_{t-1}, V_{t+1}}$; (ii) $L_{V_{t+1}|w_t, X_t^*}$; (iii) $L_{V_{t-2}, w_{t-1}, V_t}$.
3. (finite, distinctive eigenvalues) (i) for any w_t, w_{t-1}

$$0 < C_1(w_t, w_{t-1}) \leq f_{W_t|W_{t-1}, X_t^*}(w_t|w_{t-1}, x_t^*) \leq C_2(w_t, w_{t-1}) < \infty, \ \forall x_t^*$$

   (ii) for any w_t and $\overline{x}_t^* \neq \widetilde{x}_t^* \in \mathcal{X}_t^*$ there exists w_{t-1} such that

$$\frac{\partial^2 \ln f_{W_t|W_{t-1}, X_t^*}(w_t|w_{t-1}, \overline{x}_t^*)}{\partial w_t \partial w_{t-1}} \neq \frac{\partial^2 \ln f_{W_t|W_{t-1}, X_t^*}(w_t|w_{t-1}, \widetilde{x}_t^*)}{\partial w_t \partial w_{t-1}}.$$

4. (normalization) for any w_t, $x_t^* = G\left\{f_{V_{t+1}|W_t, X_t^*}(\cdot|w_t, x_t^*)\right\}$

Main results

- **Theorem 1**: Under assumptions above, the density $f_{W_{t+1},W_t,W_{t-1},W_{t-2},W_{t-3}}$ uniquely determines $f_{W_t,X_t^*|W_{t-1},X_{t-1}^*}$
- **Corollary 1**: 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

Initial conditions

- **Corollary 2** (Non-stationary case): Under assumptions above, the density $f_{W_{t+1},W_t,W_{t-1},W_{t-2},W_{t-3}}$ uniquely determines $f_{W_{t-1},X^*_{t-1}}$.
- **Corollary 3** (Stationary case): the density $f_{W_{t+1},W_t,W_{t-1},W_{t-2}}$ uniquely determines $f_{W_{t-2},X^*_{t-2}}$.

In each case, these identified objects can be used as sampling densities for initial conditions. (Two step estimation methods for dynamic models may require this.)

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 $[L, U]$:

$$X_t^* = 0.5X_{t-1}^* + 0.3\psi\left(M_{t-1}\right) + 0.2\nu_t$$

  - ν_t are i.i.d. truncated normal on $[L, U]$.
  - $\psi\left(M_{t-1}\right) = L + (U - L)\,\frac{\exp(M_{t-1})-1}{\exp(M_{t-1})+1}$,

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.	$M_{t+1} = \begin{cases} M_t + \exp(\eta_{t+1} + X_{t+1}^*) \\ \exp(\eta_{t+1} + X_{t+1}^*). \end{cases}$ ...

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

Assumption 2: invertibility assumptions

Use $V_t = M_t$ (continuous element of W_t)

- $L_{M_{t+1},w_t|w_{t-1},M_{t-2}}$: Consider w_t w/ $y_t = 1$.
  - **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}.
- $L_{M_{t+1}|w_t,X^*_t}$: Again, consider w_t w/ $y_t = 1$.
  - **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.
- Assumption 2(iii): similar argument to 2(i)
- For Spec B: appendix discusses sufficient conditions
- NB: One-to-one rules out models where W_t only has discrete components, but X^*_t is continuous.
- NB: when M_{t+1} depends just on w_t, but not on X^*_{t+1}, then cannot use $V_t = M_t$: "too little feedback".

Assumption 3: Finite, distinct eigenvalues

1. 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. 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

$$\frac{\partial^2}{\partial m_t \partial m_{t-1}} \ln f_{W_t|W_{t-1},X_t^*}(w_t|w_{t-1}, x_t^*) \quad \text{varies in } x_t^*$$

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

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

where $\tilde{\phi}(\cdot)$ is N(0,1) density truncated to [0,1].

$\frac{\partial^2}{\partial m_t \partial m_{t-1}} \ln f = \frac{\partial^2}{\partial m_t \partial m_{t-1}} \left(\log\left(\frac{m_t - m_{t-1}}{\exp(X_t^*)}\right)\right)^2$. Cdtn holds.

Spec. A: $m_t|m_{t-1}, y_{t-1}, X_t^*$ is never function of X_t^*. Cannot hold.

Assumption 4

Appropriate normalization to pin down unobserved X_t^*

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

$$h(w_t, z) = (1 - y_t)m_t + C_{med} \cdot \exp(0.3\psi(m_t)) \cdot \exp(0.5z)$$

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

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

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

$\widehat{f}(W_t, X_t^* \mid W_{t-1}, X_{t-1}^*)$	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

Extensions

1. Companion work on dynamic games
   - X_t^* is multivariate (X_t^* includes USV's for each player).
   - For example, dynamic capacity investment
     - $Y_t = (Y_{1t}, Y_{2t})$: each firm's capacity investment
     - $M_t = (M_{1t}, M_{2t})$: each firm's total capacity
     - $X_t^* = (X_{1t}^*, X_{2t}^*)$: each firm's productivity
   - Consider alternatives to LF:

$$f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*} = \underbrace{f_{Y_t | M_t, X_t^*}}_{\text{CCP}} \cdot \underbrace{f_{X_t^* | M_t, M_{t-1}, X_{t-1}^*}}_{X \text{ transition}} \cdot \underbrace{f_{M_t | Y_{t-1}, M_{t-1}, X_{t-1}^*}}_{M \text{ transition}}$$

   - Can apply arguments in Hu-Schennach (2008):

$$f_{Y_t, M_t, Y_{t-1} | M_{t-1}, Y_{t-2}} = \int f_{Y_t | M_t, M_{t-1}, X_{t-1}^*} f_{M_t, Y_{t-1} | M_{t-1}, X_{t-1}^*} f_{X_{t-1}^* | M_{t-1}, Y_{t-2}} dx_{t-1}^*$$

   - Assumption 4 more complicated: monotonicity not enough.
2. Two-step estimation (as in HM, BBL):
   - Estimate CCP, LOM by sieve MLE
   - Estimate structural parameters from optimality conditions

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: law of motion $f_{W_t, X_t^* | W_{t-1}, X_{t-1}^*}$ identified from $f_{W_{t+1}, W_t, W_{t-1}, W_{t-2}, W_{t-3}}$ (5 obs.)
  2. stationary: law of motion $f_{W_2, X_2^* | W_1, X_1^*}$ identified from $f_{W_{t+1}, W_t, W_{t-1}, W_{t-2}}$ (4 obs.)
- More broadly: apply measurement error econometrics to non-measurement error settings:
  - Auction models: (i) unobserved # bidders; (ii) unobserved heterogeneity
  - Price search models, where number of firms not observed (only prices)

Details on limited feedback

1. Learning model (Crawford-Shum; Ching; Erdem-Keane)
   - Y_t: choice of drug treatment
   - M_t: # times drug has been tried
   - X_t^*: current beliefs ("posterior mean") regarding drug effectiveness
2. Dynamic stockpiling model (Hendel-Nevo)
   - Y_t: brand of detergent purchased
   - M_t: inclusive values from each detergent brand
   - X_t^*: inventory of detergent

- In both these models, evolution of M_t depends just on (Y_{t-1}, M_{t-1}), not on X_t^* or X_{t-1}^*.
- Note: little restriction on evolution of X_{t+1}^*, can depend on $X_{t-1}^*, Y_{t-1}, M_{t-1}$.

Return

Flowchart

Return