Nonparametric Identification of Dynamic Models with Unobserved State Variables, Supplemental Material

Yingyao Hu
Johns Hopkins University

Matthew Shum
California Institute of Technology

May 2009

Abstract
We provide some additional material pertaining to our paper Hu and Shum (2008). Section 1 contains additional discussion of the identification assumptions, in the context of specific examples drawn from the empirical IO literature. Section 2 contains additional discussion of Assumption 2.

1 Additional remarks on Examples

1.1 Rust’s (1987) bus engine replacement model

In addition to the two examples presented in the main paper, we present here a discussion of our assumptions in the context of a third example: Rust’s (1987) bus-engine replacement model, augmented to allow for persistent unobserved state variables. In this model, \( W_t = (Y_t, M_t) \), where \( Y_t \) is the indicator that the bus engine was replaced in week \( t \), and \( M_t \) is the mileage since the last engine replacement.

As in the generalized investment example from the main text, we will restrict \( X_t^* \) to have a bounded support: for \([0, U]\) such that \( 0 < U < +\infty \),

\[
X_t^* = 0.5X_{t-1}^* + 0.3\psi(M_{t-1}) + 0.2\nu_t; \quad \psi(M_{t-1}) = U\frac{e^{M_{t-1}} - 1}{e^{M_{t-1}} + 1}. \tag{1}
\]

\( \nu_t \) is a truncated standard normal shock over the interval \([0, U]\), distributed independently over \( t \). We also assume that the initial value \( X_0^* \in [0, U] \), which guarantees that \( X_t^* \in [0, U] \) for all \( t \). Hence, \( X_t^*|X_{t-1}^*, Y_{t-1}, M_{t-1} \) is distributed with density determined by \( f_{\nu_t}(\cdot) \).
Let $S_t \equiv (M_t, X_t^*)$ denote the persistent state variables in this model. The period utility from each choice is additive in a function of the state variables $S_t$, and a choice-specific non-persistent preference shock:

$$u_t = \begin{cases} u_0(S_t) + \epsilon_{0t} & \text{if } Y_t = 0 \\ u_1(S_t) + \epsilon_{1t} & \text{if } Y_t = 1 \end{cases}$$

where $\epsilon_{0t}$ and $\epsilon_{1t}$ are i.i.d. Type I Extreme Value shocks, which are independent over time, and also independent of the state variables $S_t$.

The choice-specific utility functions are:

$$u_0(S_t) = -c(M_t); \quad u_1(S_t) = -RC. \quad (2)$$

In the above, $c(M_t)$ denotes the maintenance cost function, which is increasing in mileage $M_t$, and $0 < RC < +\infty$ denotes the cost of replacing the engine. We also assume that the maintenance cost function $c(\cdot)$ is bounded below and above: $c(0) = 0$; $\lim_{M \to +\infty} c(M) = \bar{c} < +\infty$. Mileage evolves as:

$$M_{t+1} - (1 - Y_t)M_t = \exp(\eta_{t+1} + X_{t+1}^*). \quad (3)$$

where $\eta_{t+1} > 0$ follows a standard normal random variable, truncated to $[0, 1]$, with density $\tilde{\phi}(\eta) \equiv \frac{\phi(\eta)}{\Phi(1) - \Phi(0)}$, where $\phi$ and $\Phi$ denote the standard normal density and CDF.\(^1\) Hence, $X_t^*$ affects the evolution of mileage, but not the agent’s utilities. Furthermore, following Rust’s assumptions, previous mileage $M_{t-1}$ has no direct effect on current mileage $M_t$ when the engine was replaced in the previous period ($Y_{t-1} = 1$).

This is a stationary dynamic optimization model, in which the conditional choice probabilities take the multinomial logit form (for $Y_t = 0, 1$): $P(Y_t|S_t) = \exp(V_y(S_t)) / \left[ \sum_{y'=0}^{1} \exp(V_{y'}(S_t)) \right]$ where $V_y(S_t)$ is the choice-specific value function in period $t$, defined recursively by $V_y(S_t) = u_y(S_t) + \beta E \left[ \log \left( \sum_{y'=0}^{1} \exp \left( V_{y'}(S_{t+1}) \right) \right) | Y_t = y, S_t \right]$. We consider each assumption in turn.

**Assumption 1** is satisfied for this model.

**Assumption 2** contains three invertibility assumptions. For the $V_t$ variables in Assumption 2, we use $V_t = M_t$, for all periods $t$. As in the generalized investment example, we begin by verifying Lemma 4 from the main text, which has a necessary condition for an operator to be one-to-one. For convenience, we reproduce that Lemma here:

\(^1\)For this to be reasonable, assume that mileage is measured in units of 10,000 miles.
Lemma 4 (Necessary conditions for one-to-one): If \( L_{R_2,R_3} \) is one-to-one, then for any set \( S \subseteq R_3 \) with \( \Pr(S) > 0 \), there exists a set \( S_1 \subseteq R_1 \) such that \( \Pr(S_1) > 0 \) and
\[
\frac{\partial}{\partial r_3} f_{R_2,R_3}(r_1,r_3) \neq 0 \text{ almost surely for } \forall r_1 \in S_1, \forall r_3 \in S.
\] (4)


We first consider Assumption 2(i). Pick any \( w_t \). Because \( X_t^* \) directly enters the mileage process, the distribution of \( M_{t+1} \) depends on \( X_{t+1}^* \). Similarly, the distribution of \( M_{t-2} \) depends on \( X_{t-2}^* \). Since \( (X_{t+1}^*,X_{t-2}^*) \) are correlated, the density of \( (M_{t+1},w_t,w_{t-1},M_{t-2}) \) varies in \( M_{t-2} \), for different values of \( (M_{t+1},w_t,w_{t-1}) \). The discussion of Assumption 2(iii) is very similar to that of 2(i), and we omit it for convenience here.

Assumption 2(ii) requires that, for all \( w_t \), the mapping \( L_{M_{t+1}|w_t,X_{t}^*} \) is one-to-one. As before, for any \( w_t \), \( M_{t+1} \) is distributed according to a mixture distribution which depends on \( X_{t+1}^* \). Since \( X_{t+1}^* \) and \( X_t^* \) are serially correlated, \( M_{t+1} \) will vary in \( X_t^* \), for fixed \( w_t \).

Here we have just shown that necessary conditions for Assumption 2 hold in this example. However, because the laws of motion (1) and (3) are both either linear or log-linear, we can also verify sufficient conditions for Assumption 2, as we did in Appendix B in the main paper, for the generalized investment example. Since the arguments are very similar to that example, we do not repeat them here.

Assumption 3 contains two restrictions on the density \( f_{W_t|W_{t-1},X_t^*} \), which factors as
\[
f_{W_t|W_{t-1},X_t^*} = f_{Y_t|M_t,X_t^*} \cdot f_{M_t|Y_{t-1},M_{t-1},X_t^*}.
\] (5)

Assumption 3(i) requires that, for any \( (w_t,w_{t-1}) \), this density is bounded between 0 and +\( \infty \). The first term is the CCP \( f_{Y_t|M_t,X_t^*} \), which is a logit probability. Because the per-period utilities, net of the \( \epsilon \)'s, are bounded away from \(-\infty \) and +\( \infty \), the logit choice probabilities are also bounded away from zero. The second term is the mileage law of motion \( f_{M_t|Y_{t-1},M_{t-1},X_t^*} \) which, by assumption, is a truncated normal distribution, so it is also bounded away from zero and +\( \infty \). The bounded support assumption on \( M_t \) is crucial but, in practice, imply little loss in generality, because typically in estimating these models, one can take the upper and lower bounds on \( M_t \) from the observed data.

Assumption 3(ii) ensures that the eigenvalues in the decomposition (Eq. (12) in the main paper) are distinctive. Because of the factorization (5), and the fact that the CCP’s
are bounded away from zero, a sufficient condition for Eq. (3) in the main paper is that

$$\frac{\partial^2}{\partial m_t \partial m_{t-1}} \ln f_{M_t|Y_{t-1}, M_{t-1}, X^*_t}(m_t|y_{t-1}, m_{t-1}, x^*_t)$$

is strictly monotonically in \(x^*_t\), for all \(m_t, x^*_t\), and some \(w_{t-1} = (y_{t-1}, m_{t-1})\).

For any value of \(m_t\), pick any \(m_{t-1}\) such that \(y_{t-1} = 0\) (i.e., the bus engine was not replaced in period \(t-1\)). The density of \(M_t|Y_{t-1}, M_{t-1}, X^*_t\) for this pair of \((m_t, m_{t-1})\), is distributed with density 

$$\tilde{\phi} \left( \log \left( \frac{m_t - m_{t-1}}{\exp(x^*_t)} \right) \right) / \left| m_t - m_{t-1} \right|$$

on the range \(m_t \in [m_{t-1}, m_{t-1} + \exp(x^*_t)]\), where \(\tilde{\phi}(\cdot)\) denotes a truncated standard normal density. The second derivative of the log of this density is monotonic in \(x^*_t\).

**Assumption 4** presumes a known functional \(G\) such that \(G[f_{M_{t+1}|Y_t, M_t, X^*_t}(\cdot|y_t, m_t, x^*_t)]\) is monotonic in \(x^*_t\). Eqs. (1) and (3) imply that

$$M_{t+1} = (1 - Y_t) M_t + \exp(\eta_{t+1} + 0.2 \nu_{t+1}) \cdot \exp(0.3 \psi (M_t)) \cdot \exp(0.5 X^*_t).$$

Let \(C_{med}\) denote the median of the random variable \(\exp(\eta_{t+1} + 0.2 \nu_{t+1})\), which is a truncated log-normal random variable. Then

$$\text{med} \left[ f_{M_{t+1}|Y_t, M_t, X^*_t}(\cdot|y_t, m_t, x^*_t) \right] = (1 - y_t) m_t + C_{med} \cdot \exp(0.3 \psi (m_t)) \cdot \exp(0.5 x^*_t)$$

which is monotonic in \(x^*_t\). Hence, we can pin down \(x^*_t = \text{med} \left[ f_{M_{t+1}|Y_t, M_t, X^*_t}(\cdot|y_t, m_t, x^*_t) \right]\).

### 1.2 Remark on investment models

For Example 2 in the main paper, we considered a general investment model in the framework of Doraszelski and Pakes (2007). There is a recent and growing empirical literature based on these types of dynamic models, including Collard-Wexler (2006), Ryan (2006), and Dunne, Klimer, Roberts, and Xu (2006). Pakes (2008, section 3) and Ackerberg, Benkard, Berry, and Pakes (2007) discuss additional examples.

On the other hand, the productivity literature has by and large been based on the "pure" investment model, typified by Olley and Pakes (1996) (OP). This model differs in an important way from the types of models considered in our paper. Namely, in OP, capital stock (corresponding to the \(M\) variable in Example 2) evolves deterministically, conditional on the previous period’s capital \((M_{t-1})\) and investment \((Y_{t-1})\). This feature violates two of our maintained assumptions (\# 2,3), which require that \(M_t\) depend on \(X^*_t\) even conditional
on \((Y_{t-1}, M_{t-1})\). For this reason, in Example 2 in the main paper, we do not consider the “pure” investment model as in OP, but rather a generalized investment model in which \(M_t\) does not evolve deterministically.

2 Further discussion on Assumption 2

In this section we discuss how Assumption 2 is used to ensure the validity of two different ways for taking operator inverses. Consider two scenarios involving an operator equation

\[ L_{R_1,r_2,R_4} = L_{R_1| r_2,R_3} L_{r_2,R_3,R_4}. \]  

(8)

In the first scenario, suppose we want to solve for \(L_{r_2,R_3,R_4}\) given \(L_{R_1| r_2,R_3}\) and \(L_{R_1,r_2,R_4}\). The assumption that \(L_{R_1| r_2,R_3}\) is one-to-one guarantees that we may have

\[ L_{R_1| r_2,R_3}^{-1} L_{R_1,r_2,R_4} = L_{r_2,R_3,R_4}. \]  

(9)

As an example, Assumption 2(ii) guarantees that pre-multiplication by the inverse operator \(L_{R_1| w_t, X_t}^{-1}\) is valid, which is used in the equation following Eq. (9).

In the second scenario, suppose we need to solve for \(L_{R_1| r_2,R_3}\) given \(L_{R_1,r_2,R_4}\) and \(L_{r_2,R_3,R_4}\) in equation (8). We would need the operator \(L_{r_2,R_3,R_4}\) to be invertible as follows:

\[ L_{R_1,r_2,R_4} L_{r_2,R_3,R_4}^{-1} = L_{R_1| r_2,R_3}. \]  

(10)

As proved in Lemma 1 in Hu and Schennach (2008), the sufficient condition for obtaining Eq. (10) from Eq. (8) is that the operator \(L_{R_4,R_3,r_2}\) is one-to-one.\(^2\) (Notice that the operator \(L_{R_4,R_3,r_2}\) is from \(L^p(\mathcal{R}_3)\) to \(L^p(\mathcal{R}_4)\).)

Assumption 2(i) is an example of this. It is used to justify the post-multiplication by \(L_{V_{t+1}| w_{t-1}, V_{t-2}}^{-1}\) and \(L_{V_{t+1}, w_t| w_{t-1}, V_{t-2}}^{-1}\) in, respectively, Eqs. (10) and (11). Similarly, Assumption 2(iii) guarantees that post-multiplication by \(L_{V_t| w_{t-1}, V_{t-2}}^{-1}\), which is used in the second line of the bottom display on pg. 21. Throughout this paper, we only post-multiply by the inverses of \(L_{V_{t+1}, w_t| w_{t-1}, V_{t-2}}\) and \(L_{V_t| w_{t-1}, V_{t-2}}\); all other cases of inverses involve pre-multiplication. For a more technical discussion, see Aubin (2000, sections 4.5-4.6).

\(^2\) A similar assumption is also used in Carroll, Chen, and Hu (2009).
References


