Sequential Estimation in Discrete Decision Problems:

CCP and Nested Pseudo Likelihood

Dynamic Programming and Structural Econometrics #8

Bertel Schjerning, University of Copenhagen

Do we need to solve the DP problem to estimate model?

Structural estimation of DP problems is hard

- NFXP can be computationally expensive since we need to repeatedly compute EV or V_σ as fixed points to the Bellman operator.
- What if we could estimate them from the data - without solving the nested fixed point problem?

Pathbreaking paper:

Hotz and Miller (ReStud, 1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models"

Hotz-Miller Inversion

- Hotz and Miller's idea is to use observable data on x and d to estimate of $P(d|x)$ and then by their inversion theorem map $P(d|x)$ on to (differences in) the value function.

CCP estimators

We can then use the two step estimator

1. Estimate reduced form conditional choice probabilities (CCPs) using data on x and d. Label them $\hat{P}(d|x)$
2. Use the Hotz-miller inversion to map $\hat{P}(d|x)$ to estimated value function differences and thereby "measure" continuation values that enter in the sample criterion (e.g. likelihood).

We refer to this as the Hotz-Miller approach or the CCP estimator

The power of the Hotz-Miller inversion

We use the Inversion Theorem to:

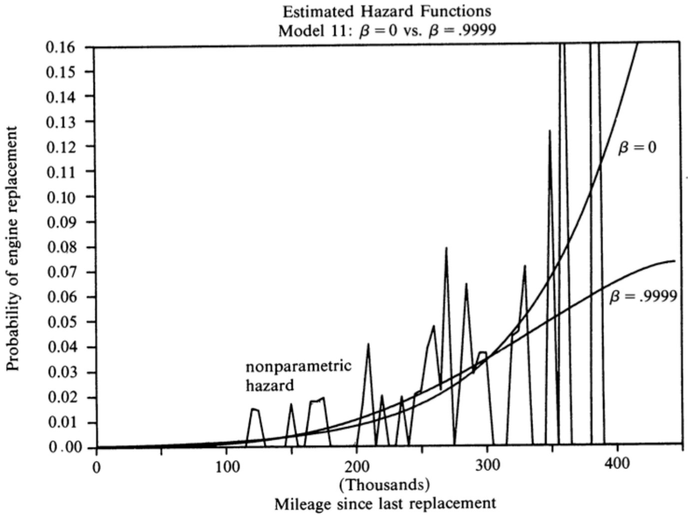
1. Develop two-step estimators to estimate structural parameters, θ without solving the model.
2. Provide empirically tractable representations of the conditional value functions.
3. Analyze identification in dynamic discrete choice models.
4. Introduce new methods for incorporating unobserved heterogeneity using the EM algorithm
   ....see Arcidiacono and Miller (ECMA, 2011)
5. Exploit finite dependence when estimating non-stationary models.

Although consistent, two-step CCP estimators are inefficient and biased in small samples.

- May be hard to obtain Non-Parametric estimates of $P(d|x)$
- In small samples Bellman equation does not necessarily hold.

Small sample problems

Sometimes it can be hard to get a precise nonparametric estimate of CCP



Sequential Estimation

Aguirregabiria and Mira (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models"

Nested Pseudo Likelihood (NPL) algorithm

- Use Hotz-Miller inversion to express solution of the DP problem in choice probability space (rather than value functions space)
- Recursively update $\hat{P}(d|x)$ to obtain a sequence of estimators: $\hat{\alpha}_k$ with $k = 1 \ldots, K$

Statistical and computational properties of the estimator

- When NPL is initialized with consistent nonparametric estimates of CCPs the sequence includes as extreme cases
  1. The Hotz-Miller CCP estimator (for K = 1)
  2. Rust's nested fixed point MLE estimator (in the limit when $K \to \infty$).
- Trade-off between finite sample precision and computational cost in the sequence of policy iteration estimators.
- Monte Carlo: based on Rust's bus replacement model

The General Problem

Bellman equation

$$V(s;\theta) = \max_{a \in \mathcal{A}(s)} \{u(s,a;\theta_u) + \beta \int V(s';\theta) p(s'|s,a;\theta_g,\theta_f) ds'\}$$

u and p: known up to a set of parameters, $\theta = (\theta_u, \theta_g, \theta_f)$

- **The agent's problem:** Maximize expected sum of current and future discounted utilities
  - a : Discrete control variable, $a \in \mathcal{A}(s) = \{1, 2, .., J\}$.
  - s : Current state, fully observed by agent
  - $s\prime$: Future state; possibly continuous and subject to uncertainty
- **The agents beliefs about s′:**
  - Obeys a (controlled) Markov transition probability $p(s_{t+1}|s_t, a_t; \theta_g, \theta_f)$
- **Model solution,** $V(s;\theta)$
  - Find the fixed point for the Bellman equation

A&M maintain Rust's Assumptions

Assumption (Conditional Independence (CI))

State variables, $s_t = (x_t, \varepsilon_t)$ obeys a (conditional independent) controlled Markov process with probability density

$$p(x_{t+1}, \varepsilon_{t+1} | x_t, \varepsilon_t, a, \theta_g, \theta_f) = g(\varepsilon_{t+1} | x_{t+1}, \theta_g) f(x_{t+1} | x_t, a, \theta_f)$$

Assumption (Additive Separability (AS))

$$U(s_t, a) = u(x_t, a; \theta_u) + \varepsilon_t(a)$$

Assumption (Finite Domain of Observable State Variables)

$$x \in X = \{x^1, x^1, ..., x^m\}$$

Assumption (XV)

The unobserved state variables, ε_t are assumed to be multivariate iid. extreme value distributed

Bellman equation and choice probabilities

Define the smoothed value function $V_\sigma(x) = \int V(x,\epsilon)g(\epsilon|x)d\epsilon$ where σ represents parameters that index the distribution of the $\epsilon's$.
($\sigma = \theta_2$ in Rust notation)

Under assumptions CI, AS and finite domain of x, we can summarize the solution by the smoothed Bellman operator, $\Gamma_\sigma(V_\sigma)$

$$V_\sigma(x) = \int \max_{a \in \mathcal{A}(x)} \left\{ u(x,a) + \epsilon(a) + \beta \sum_{x'} V_\sigma(x') f(x'|x,a) \right\} g(\epsilon|x) d\epsilon$$

The conditional choice probability (CCP)

$$P(a|x) = \int I\{a = \arg\max_{j \in \mathcal{A}(x)} \{v(j,x) + \epsilon(j)\}\, g(\epsilon|x) d\epsilon$$

where $v(x,a) = u(x,a) + \beta \sum_{x'} V_\sigma(x') f(x'|x,a)$ is the choice-specific value function

From Conditional Choice Probabilities to Value Functions

- $P(a|x)$ is uniquely determined by the vector of normalized value function differences $\tilde{v}(x,a) = v(x,a) - v(x,1)$
- That is, there exists a vector mapping $Q_x(.)$ such that $\{P(a|x) : a > 1\} = Q_x(\tilde{v}(x,a) : a > 1)$, where, without loss of generality, we exclude the probability of alternative one.
- Hotz and Miller establish that this mapping is invertible under CI.
- In general

$$Q_x^j(\tilde{v}(x,a) : a > 1) = \partial S([0, \{\tilde{v}(x,a) : a > 1\}], x)/\partial \tilde{v}(x,j)$$

where

$$S(\{v(x,a) : a \in A\}, x) = \int \max[v(x,a) + \epsilon(a)] g(d\epsilon|x)$$

is McFadden's social surplus function.

From Conditional Choice Probabilities to Value Functions

- Under assumption (XV), social surplus function is the well known "log-sum" formula

$$
\begin{aligned}
S(\{v(x,a) : a \in A\}, x) &= \int \max_{a \in A}[v(x,a) + \epsilon(a)]g(d\epsilon|x) \\
&= \sigma \log \sum_{j \in A} exp(v(x,j)/\sigma)
\end{aligned}
$$

  the j'th component Q_x takes the well known logistic form

$$
Q_x^j(\tilde{v}(x,a)) = \frac{\exp(\tilde{v}(x,a)/\sigma)}{1 + \sum_{j=2}^{A} \exp(\tilde{v}(x,j)/\sigma)}
$$

- NOTE, it's not hard to invert Q_x in this case

From Conditional Choice Probabilities to Value Functions

The Smooth Bellman equation can be re-written as

$$V_\sigma(x) = \sum_{a \in A} P(a|x) \left\{ u(x,a) + E[\epsilon(a)|x,a] + \beta \sum_{x'} V_\sigma(x') f(x'|x,a) \right\}$$

where $E[\epsilon(a)|x,a]$ is the expectation of the unobservable ϵ conditional on the optimal choice of alternative a:

$$\begin{aligned} E[\epsilon(a)|x,a] = [P(a|x)]^{-1} \int \epsilon(a) I\{\tilde{v}(x,a) + \epsilon(a) \\ > \tilde{v}(x,k) + \epsilon(j), j \in A(x)\} g(d\epsilon|x) \end{aligned}$$

$E[\epsilon(a)|x,a]$ clearly depends on $\tilde{v}(x,a)$, but due to the invertibility of Q_x we can express it probability space

$$E[\epsilon(a)|x,a] = e_x(a, \{P(j|x)\}_{j \in A}).$$

Under XV we have $E[\epsilon(a)|x,a] = \gamma - \ln P(a|x)$ where $\gamma = 0.5772156649...$ is Euler's constant

From Conditional Choice Probabilities to Value Functions

In compact matrix notation we can write

$$V_\sigma = \sum_{a \in A} P(a) * \{u(a) + e(a, P) + \beta F(a) V_\sigma\}$$

where $*$ is the element by element product and $P(a)$, $u(a)$, $e(a, P)$ and V_σ are all $M \times 1$ vectors and $F(a)$ is the $M \times M$ matrix of conditional transition probabilities $f(x'|x, a)$

This system of fixed point equations can be solved for the value function to obtain V_σ as a function of P:

$$V_\sigma = \psi(P) = [I - \beta F^U(P)]^{-1} \sum_{a \in A} \{P(a) * [u(a) + e(a, P)]\}$$

"policy valuation operator" = mapping from choice probabilities to smooth value functions

where $F^U(P) = \sum_{a \in A} P(a) F(a)$ is the $M \times M$ matrix of unconditional transition probabilities induced by P.

The Fixed Point Problem in Probability Space

Recall that

$$V_\sigma = \psi(P) = [I - \beta F^U(P)]^{-1} \sum_{a \in A} \{P(a) * [u(a) + e(a, P)]\} \qquad (1)$$

and

$$P(a|x) = \int I\{a = \arg\max_{j \in \mathcal{A}(x)} \{v(j, x) + \epsilon(j)\} \, g(\epsilon|x) d\epsilon \qquad (2)$$

where $v(x, a) = u(x, a) + \beta \sum_{x'} V_\sigma(x') f(x'|x, a)$

The policy iteration operator, Ψ

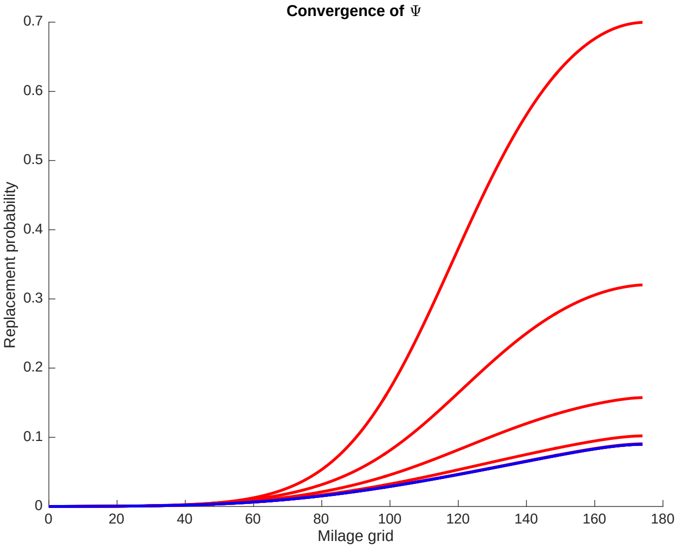
Substituting the *policy valuation operator*, $\psi(P)$ defined by (1) into the formula for CCP's in (2) we obtain the cornerstone of NPL algorithm algorithm:

$$P = \Psi(P) = \Lambda(\psi(P))$$

- Hence, the optimal choice probabilities P is a fixed point of Ψ.
- Thus the original fixed point problem in "value space" can be reformulated as a fixed point problem in "probability space"

Finding fixed point, $P = \Psi(P)$

Fast convergence of successive approximations, $P_{k+1} = \Psi(P_k)$



Likelihood function

Data: $(a_{i,t}, x_{i,t})$, $t = 1, ..., T_i$ and $i = 1, ..., n$

Likelihood function

$$\ell_i^f(\theta) = \ell_i^1(\theta) + \ell_i^2(\theta_f) = \sum_{t=2}^{T_i} log(P_\theta(a_{i,t}|x_{i,t})) + \sum_{t=2}^{T_i} log\left(f_{\theta_f}(x_{i,t}|x_{i,t-1}, a_{i,t-1})\right)$$

Two Step-Estimator

- Consistent estimates of the conditional transition probability parameters θ_f can be obtained from transition data without having to solve the Markov decision model.
- We focus on the estimation of $\alpha = (\theta_u, \theta_g)$ given initial consistent estimates of θ_f obtained from maximizing the partial log-likelihood $\ell^2(\theta_f) = \sum_i \ell_i^2(\theta_f)$
- Originally suggested in Rust(1987)

Nested Pseudo Likelihood Algorithm

Initialization

- Let $\hat{\theta}_f$ be an estimate of θ_f .
- Start with an initial guess for the conditional choice probabilities, $P^0 \in [0,1]^{MJ}$.

At iteration $K \geq 1$, apply the following steps:

- Step 1: Obtain a new pseudo-likelihood estimate of α, α^K, as

$$\alpha^K = \arg\max_{\alpha \in \Theta} \sum_{i=1}^{n} \ln \Psi_{\alpha,\hat{\theta}_f}(P^{K-1})(a_i|x_i) \tag{3}$$

  where $\Psi_\theta(P)(a|x)$ is the (a,x)'s element of $\Psi_\theta(P)$.
- Step 2: Update P using the arg max from step 1, i.e.

$$P^K = \Psi_{(\alpha^K,\hat{\theta}_f)}(P^{K-1}) \tag{4}$$

- Iterate in K until convergence in P (and α) is reached.

Sequential Policy Iteration Estimators

The K-stage PI estimator is defined as:

$$\hat{\alpha}^K = \arg\max_{\alpha \in \Theta} \sum_{i=1}^{n} \ln \Psi(P^{K-1})(a_i|x_i)$$

where $P^K = \Psi_{(\hat{\alpha}^K, \hat{\theta}_f)}(P^{K-1})$ and P^0 is a consistent, non-parametric estimate of the true conditional choice probabilities

- Performing one, two, and in general K iterations of the NPL algorithm yields a sequence $\{\hat{\alpha}_1, \hat{\alpha}_1, ..., \hat{\alpha}_K\}$ of statistics that can be used as estimators of the true value of α, α^*
- A&M call them sequential policy iteration (PI) estimators.

Statistical properties of K-PI estimator

For any K

- $\hat{\alpha}^K$ is asymptotically equivalent to MLE
- $\hat{\alpha}^K$ is $\sqrt{n}$ consistent
- $\hat{\alpha}^K$ is asymptotic normal with known variance-covariance matrix (A&M has an expression that accounts for first step estimation error)

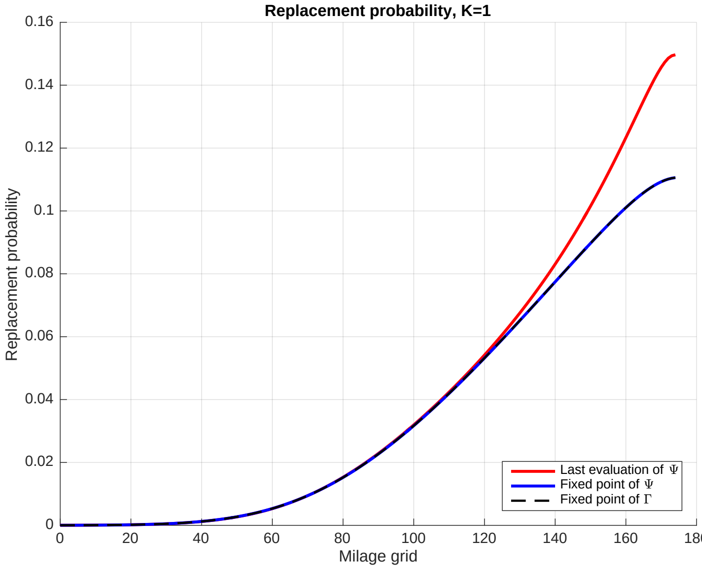
For $K = 1$

- $\hat{\alpha}^K$ encompasses Hotz-Miller (1993) estimator

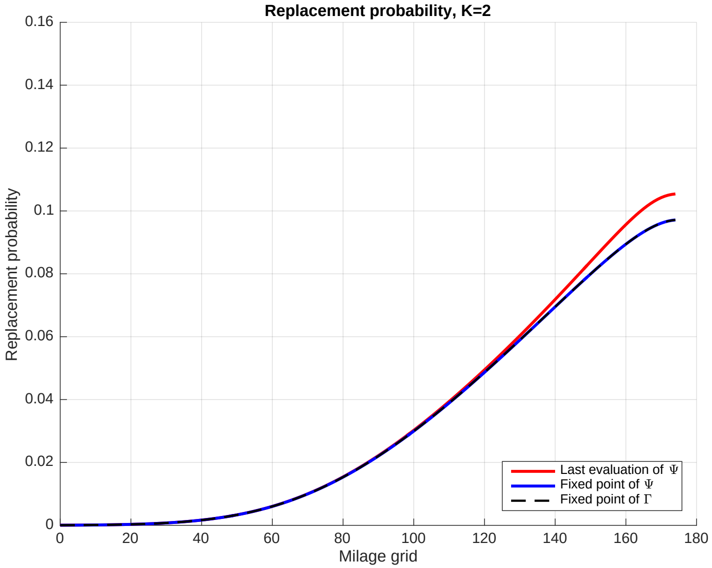
As $K \to \infty$

- $\hat{\alpha}^K$ converges to the MLE estimator obtained by NFXP
- Standard inference.

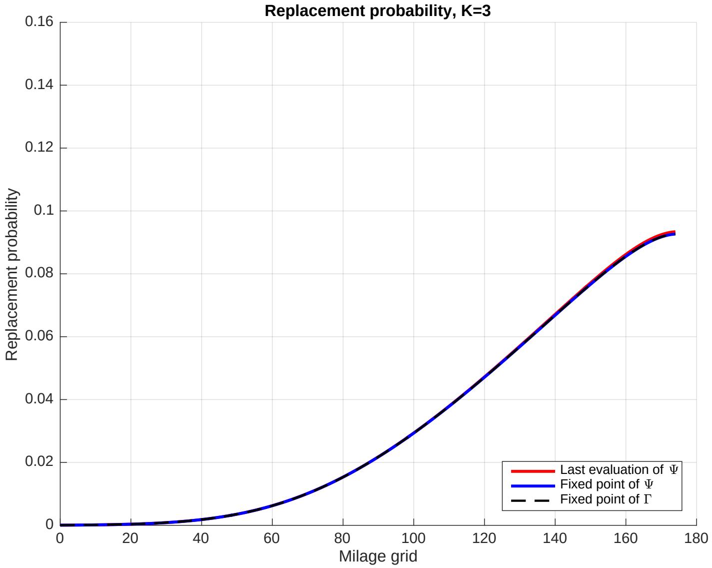
Replacement probability, K=1



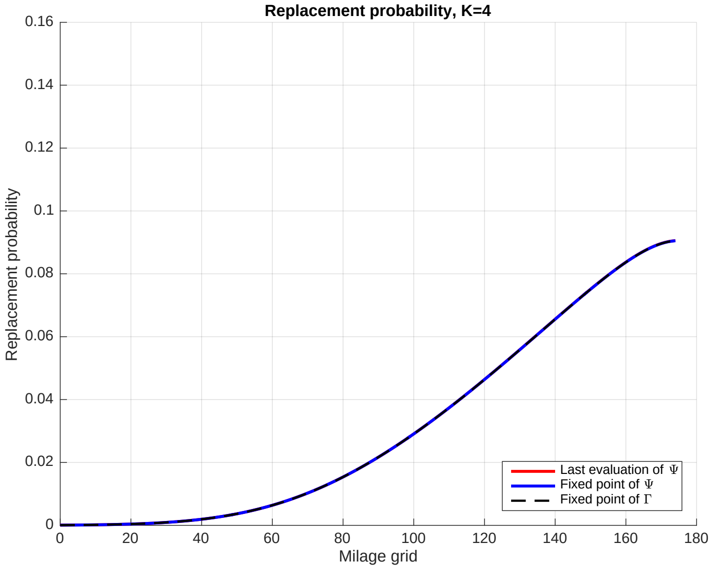
Replacement probability, K=2



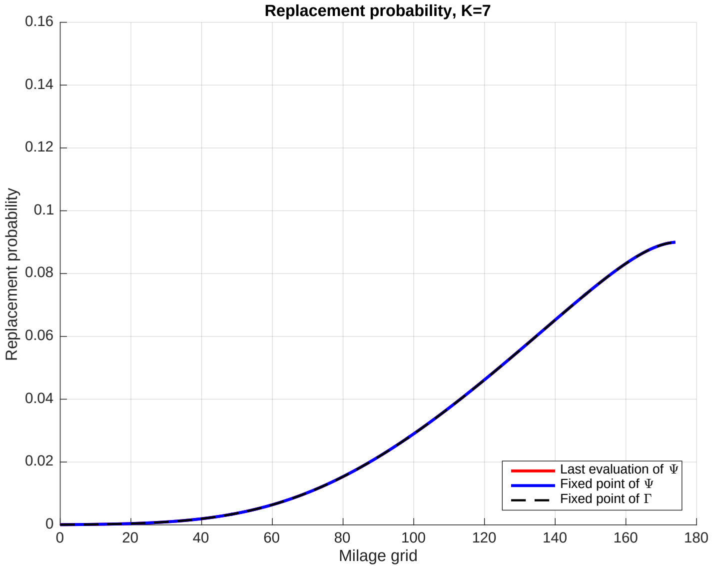
Replacement probability, K=3



Replacement probability, K=4



Replacement probability, MLE



Hotz-Miller's two step estimator

- The CCP estimators were defined as the values of α that solve systems of equations of the form

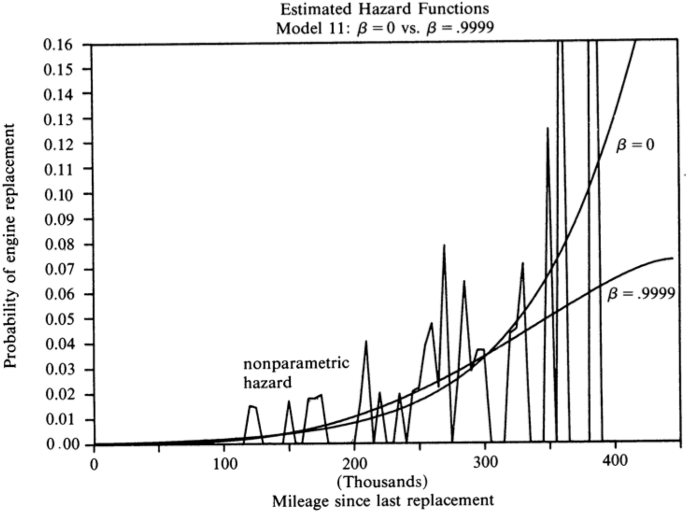
$$\arg\min_{\alpha\in\Theta}\sum_{i=1}^{N}\sum_{j=1}^{J} Z_i^j \left[I(a_i = j) - \tilde{P}_{(\alpha,\hat{\theta}_f)}(P^0)(j|x) \right]$$

  where Z_i is are vectors of instrumental variables (e.g.) functions of x_i

- Easy to show that the 1-stage PI estimator is a CCP estimator with $Z_i = \partial\Psi(P^0)(a_i|x_i)/\partial\alpha$ is used as instrument.

Small sample problems

Sometimes it can be hard to get a precise nonparametric estimate of CCP



The Precision of PI Estimators: A Monte Carlo Evidence

TABLE I

MONTE CARLO EXPERIMENT

	Experiment design
Model:	Bus engine replacement model (Rust)
Parameters:	$\theta_0 = 10.47$; $\theta_1 = 0.58$; $\beta = 0.9999$
State space:	201 cells
Number observations:	1000
Number replications:	1000
Initial probabilities:	Kernel estimates

Monte Carlo distribution of MLE
(In parenthesis, percentages over the true value of the parameter)

	θ_0	θ_1
Mean Absolute Error:	2.08 (19.9%)	0.17 (29.0%)
Median Absolute Error:	1.56 (14.9%)	0.13 (22.7%)
Std. dev. estimator:	2.24 (21.4%)	0.16 (26.9%)
Policy iterations (avg.):	6.2	

Monte Carlo distribution of PI estimators (relative to MLE)
(All entries are 100* (K-PI statistic-MLE statistic)/MLE statistic)

		Estimators		
Parameter	Statistics	1-PI	2-PI	3-PI
θ_0	Mean AE	4.7%	1.6%	0.3%
	Median AE	14.2%	0.2%	−0.3%
	Std. dev.	6.8%	1.2%	0.3%
θ_1	Mean AE	18.7%	1.5%	0.2%
	Median AE	25.1%	0.7%	0.6%
	Std. dev.	11.0%	1.3%	0.2%

The Precision of PI Estimators: A Monte Carlo Evidence

TABLE II

RATIO BETWEEN ESTIMATED STANDARD ERRORS AND STANDARD DEVIATION OF MONTE CARLO DISTRIBUTION

Parameters	Statistics	Estimators			
		1-PI	2-PI	3-PI	MLE
θ_0	Ratio	0.801	1.008	1.027	1.022
θ_1	Ratio	0.666	1.043	1.076	1.065

Relative Speed of NPL and NFXP

- For most problems the fixed point iterations (i.e., policy iterations) are much more expensive than likelihood and pseudo-likelihood "hill" climbing iterations.
- The size of the state space does not affect the number of policy iterations in any of the two algorithms.
- Both algorithms were initialized with Hotz-Miller Estimates.
- A&M found that NPL around 5 and 10 times faster than NFXP