

Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models

Victor Aguirregabiria[✉]
University of Chicago

Pedro Mira[✉]
CEMFI, Madrid

March 2000

Abstract

This paper proposes a procedure for the estimation of discrete Markov decision models and studies its statistical and computational properties. Our Nested Pseudo-Likelihood method (NPL) is similar to Rust's Nested Fixed Point algorithm (NFXP), but the nesting of the two algorithms is swapped. First, we prove that on convergence NPL produces a root of the likelihood equations. Our procedure requires fewer policy iterations at the expense of more likelihood-climbing iterations. We focus on a class of infinite-horizon, partial likelihood problems for which NPL can deliver large computational gains. Second, based on this algorithm we define a class of consistent and asymptotically equivalent Sequential Policy Iteration (PI) estimators, which encompasses both Hotz-Miller's CCP estimator and the partial Maximum Likelihood estimator. This presents the researcher with a "menu" of sequential estimators reflecting a trade-off between finite-sample precision and computational cost. Using actual and simulated data we compare the relative performance of these estimators. In all our experiments the benefits in terms of precision of using a 2-stage PI estimator instead of 1-stage (i.e., Hotz-Miller) are very significant. More interestingly, the benefits of MLE relative to 2-stage PI are small.

KEYWORDS: Discrete Markov decision models, Nested algorithms, Policy iteration, Finite sample properties.

* We are grateful to Miguel Delgado, Bob Miller, Ariel Pakes, John Rust, seminar participants at Universidad Carlos III, Carnegie Mellon University, University of Chicago and three anonymous referees for their useful comments.

1 Introduction

The computational cost of estimating discrete choice dynamic programming models remains an important constraint which has limited their range of applications.¹ There are at least two features that make the estimation of these models substantially more costly than just solving the corresponding dynamic programming problem once. First, observable and unobservable time-invariant individual heterogeneity in preferences or technology imply that the model has to be solved for each type of individual in the sample (e.g., for each combination of sex, cohort, region, family background, etc.). And second, the nested solution-estimation algorithms that are used to obtain the maximum likelihood estimator require one to solve the dynamic decision problems of all types of individuals as many times as the number of iterations needed in the search for parameter estimates. This computational constraint has substantial implications for empirical work. The model specification has to be very parsimonious in terms of state variables and time-invariant explanatory variables. In some contexts, this can lead to important misspecifications.

Recent research has resulted in some important developments in the techniques for the solution of dynamic programming models. Rust (1997a, 1997b) shows that randomization and low discrepancy methods in the discretization of the state space can break the curse of dimensionality in the solution of these models. In other related work, Keane and Wolpin (1994, 1996) recently developed a method which combines interpolation and simulation techniques for the solution and estimation of discrete choice dynamic programming models. Pakes and McGuire (1997) propose a stochastic algorithm for the computation of Markov Perfect equilibria that exploits stochastic approximation and rules used in the reinforcement learning literature. The emphasis of our paper is not in the algorithms for the solution of dynamic programs, but in the procedure that searches for the maximum likelihood estimator and in the derivation of asymptotically equivalent estimators which are computationally cheaper.

The contribution of our paper is twofold. First, we propose a new estimation

¹See Eckstein and Wolpin (1989), Pakes (1994), Rust (1994), and Miller (1997) for excellent surveys on the estimation of these models and their empirical applications.

procedure which we call Nested Pseudo-Likelihood algorithm (NPL). Our procedure is in the spirit of Rust's (1987, 1988) Nested Fixed Point algorithm (NFXP), but it is different in that the nesting of the two algorithms (likelihood climbing and fixed point) is swapped. We show that on convergence this algorithm produces a root of the likelihood equations under the same conditions as NFXP. However, our nested algorithm does not require repeated solution of the dynamic programming problems. Therefore, as the dimension of the state space and the cost of solving the dynamic programming problem increases, our nested algorithm may produce ML estimates at a smaller computational cost than NFXP.

Second, we show that when non-parametric estimates of conditional choice probabilities are used as initial guesses of our NPL algorithm, this procedure yields a sequence of consistent and asymptotically equivalent estimators. In particular, in a partial likelihood context Hotz and Miller's CCP estimator is obtained as a by-product of the first iteration. Based on this result we define a class of Policy Iteration (PI) estimators which encompasses both Hotz & Miller's and MLE. The advantage of our nested pseudo-likelihood algorithm is therefore not just a reduction of the computational cost of maximum likelihood estimation. It also presents the researcher with a 'menu' of sequential estimators reflecting a trade-off between computational cost and precision in finite samples.

Our paper builds on and extends previous results by Hotz and Miller (1993) and Manski (1991, 1993). These authors proposed and implemented estimation methods that do not require the explicit solution of the dynamic programming model. Hotz et al. (1994) combined this method with a simulation-based approach to obtain an estimator which is computationally more efficient. However, a well known important limitation of these estimators is their finite sample properties. Because they do not fully exploit the structure of the model, relying instead on non-parametric estimates of conditional choice probabilities, they may be very imprecise even under the relatively large sample sizes which are common in micro applications.² We extend Hotz and

²This limitation was pointed out by Eckstein and Wolpin (1989) and Rust (1994a, 1994b), and it has been illustrated in the context of a Monte Carlo study by Hotz et al (1994).

Miller's results by showing that: (1) their estimator is asymptotically equivalent to partial maximum likelihood; and (2) their procedure can be applied recursively to obtain estimators which have better finite sample properties than theirs but are computationally cheaper than maximum likelihood.

In Rust's NFXP algorithm the outer algorithm is a "hill climbing" procedure that maximizes the likelihood function. At each iteration of this outer procedure there is an inner fixed point algorithm that solves the dynamic programming model given the current value of the parameters. In our nested method the outer algorithm solves the fixed point problem using a policy iteration method. Each policy iteration is performed with updated parameter values which are obtained from the inner procedure. This consists of a hill climbing algorithm which maximizes the current pseudo-likelihood function where the discrete choice probabilities depend on a pseudo-value function. However, as the outer algorithm converges this pseudo-value function converges to the true value function and the pseudo-likelihood function to the true likelihood function. Our NPL algorithm reduces the number of policy iterations at the expense of a larger number of hill-climbing iterations. The reason why our algorithm may be faster than NFXP is thus quite intuitive as long as policy iterations are computationally more expensive than pseudo-likelihood climbing iterations. This is certainly the case for the example that we present in section 5, i.e., partial likelihood estimation of a class of models with extreme-value unobservables introduced by Rust (1987, 1994).

Regarding the family of K-stage Policy Iteration estimators, we analyze the trade-off[®] between finite sample precision and computational cost with several Monte Carlo experiments based on Rust's bus replacement model. We find that the benefits of using the 2-stage PI estimator instead of 1-stage (i.e., Hotz-Miller) are significant. A more interesting result is that, in all our experiments, the benefits of using MLE instead of 2-stage PI are very small. This is the case even when we generate artificially very imprecise initial guesses for the conditional choice probabilities.

The rest of the paper is organized as follows. In Section 2 we review the notation and solution methods for discrete Markov decision processes and we define a

Policy Iteration operator which is the cornerstone of our algorithm. In Section 3 we briefly review Rust's NFXP algorithm, present our alternative nested procedure, and show that it yields the MLE. In Section 4 we introduce the class of Policy Iteration estimators and we obtain their asymptotic properties. In Section 5 we compare the performance of the NFXP and NPL algorithms using the bus engine data set in Rust (1987), and present a Monte Carlo study which illustrates the precision in finite samples of the different Policy Iteration estimators. We conclude in Section 6 with a summary of our results. Proofs of Propositions and a detailed description of the algorithms are provided in the Appendix.

2 Discrete Markov Decision Processes

2.1 Definitions and notation

We define a Discrete Decision Process (DDP) and we review some assumptions and results in the literature on the estimation of DDP's. There are two types of variables in these models: the vector of state variables s and a control variable d that belongs to a finite set of mutually exclusive choice alternatives $D = \{1; 2; \dots; J\}$. Time is discrete and it is indexed by t . At each period t an agent observes s_t and chooses d_t in order to maximize the expected sum of current and future discounted utilities. Future values of some state variables are uncertain for the agent. His beliefs about uncertain future states can be represented by a Markov transition probability $p(s_{t+1}|s_t; d_t)$. The time horizon of the decision problem is infinite. Utility is time separable and $u(s_t; d_t)$ represents the one-period utility function. The parameter $\beta \in (0; 1)$ is the time discount factor.

Under some regularity conditions, Blackwell's theorem establishes the following properties. First, there exists a stationary, Markovian optimal decision rule $\pi^*(s_t)$: the decision at period t is the same as the decision at period $t + j$ if $s_t = s_{t+j}$. Therefore, we omit the subindex t for the rest of this section and we use s^0 to denote the vector of next period's state variables. Second, the value function $V(\cdot)$ is the unique solution

of the Bellman equation

$$V(s) = \max_{d \in D} \{ u(s; d) + \beta \sum_{s^0} V(s^0) p(s^0 | s; d) \} \quad (1)$$

Furthermore, let $V(s; d)$ be the value function conditional on the (hypothetical) choice of alternative d ; then $V(s; d) = u(s; d) + \beta \sum_{s^0} V(s^0) p(s^0 | s; d)$ and $s^*(s) = \arg \max_{d \in D} V(s; d)$.³

We assume that the researcher knows the utility and the transition probability functions up to a vector of parameters μ . From an econometric point of view we distinguish two types of state variables: $s = (x; \theta)$. The subvector x groups variables which are observed by both the agent and the researcher, whereas the subvector θ includes those state variables which are observed only by the agent. Given data on the values of observable state variables and the actual choices made by agents, our goal is to obtain an estimate of μ . We now introduce two assumptions concerning the role of observable and unobservable state variables which have been widely used in the literature (see Rust, 1994). These assumptions contribute to simplify the estimation problem considerably.⁴ Although the primitives and operators introduced here are functions of the parameter vector μ , in this section we are concerned with their properties for a given μ ; we therefore omit it as an argument in order to simplify the notation.

ASSUMPTION 1 (Additivity): The one period utility function is additively separable in the observable and unobservable components.

$$u(s; d) = u(x; d) + \theta(d)$$

where $\theta(d)$ is the d -th component of the $J \times 1$ vector θ . The support of $\theta(d)$ is the real line for all d .

ASSUMPTION 2 (Conditional independence): The transition probability of the

³In this paper the current period utility is not bounded. This requires a generalization of Blackwell's Theorem. See Rust (1988).

⁴From an econometric point of view, the conditional independence assumption is the stronger one. Much recent literature on the estimation of these models has been concentrated on weakening this assumption (see Pakes, 1994).

state variables factors as:

$$p(x^0; \theta^0; j; x; d; \theta) = g(\theta^0; j; x^0) \cdot f(x^0; j; x; d)$$

where $g(\cdot)$ has finite first moments and is continuous and twice differentiable in θ^0 .

ASSUMPTION 3: Finite domain for the observable state variables: $x \in X = \{x^1; \dots; x^M\}$.

We now review versions of the value functions and the Bellman operator which take into account the existence of unobservables. These versions will prove more useful than equation (1) in the analysis of the estimation problem. Recall McFadden's (1973) social surplus $S^s(x) = \int_{\mathcal{D}} V(x; \theta) g(d; j; x)$. The social surplus is the expectation of the value function conditional on the state variables x observable to the econometrician. It solves a smoothed Bellman equation which under Assumptions 1-3 takes the following form:

$$S^s(x) = \int_{\mathcal{D}} \max_{d \in \mathcal{D}} \{ u(x; d) + \theta(d) + \int_{\mathcal{X}^0} p(x^0; \theta^0; j; x; d) S^s(x^0) g(d; j; x) \} \quad (2)$$

It can be shown that the operator on the right hand side of this functional equation is a contraction mapping. Therefore, $S^s(\cdot)$ is the unique solution.⁵ From this solution we obtain the set of conditional choice probabilities. Given the vector of observable state variables x , the probability that alternative d is the optimal choice is:

$$P(d; x) = I \left(d = \arg \max_{j \in \mathcal{D}} \left\{ u(x; j) + \theta(j) + \int_{\mathcal{X}^0} p(x^0; \theta^0; j; x; d) S^s(x^0) g(d; j; x) \right\} \right) \quad (3)$$

where $I(\cdot)$ is the indicator function. Under Assumption 1 the probabilities $P(d; x)$ are strictly positive for any combination of observables $(d; x)$.

Finally, we restate in our framework an important result in Hotz and Miller (1993) which shows that conditional choice probabilities can be used to obtain an alternative representation of the social surpluses. The Bellman equation (2) can be rewritten as

$$S^s(x) = \int_{\mathcal{D}} P(d; x) \left\{ u(x; d) + E[\theta(d); j; x; \theta(x; \theta) = d] + \int_{\mathcal{X}^0} p(x^0; \theta^0; j; x; d) S^s(x^0) \right\} \quad (4)$$

⁵See the proof of (i) in Proposition 1 in the Appendix:

In multinomial choice models with Assumption 1 (separability), conditional choice probabilities and conditional expectations $E[\mathbb{1}(d=j) | x; \pm(x; \cdot) = d]$ are functions of choice-specific utility differences $v(x; j) - v(x; 1)$. Under Assumptions 1-2, $v(x; d) = u(x; d) + \sum_{x^0} \mathbf{P} f(x^0 | x; d) S^\pi(x^0)$. Since the mapping from utility differences to choice probabilities is invertible, it follows that the conditional expectations $E[\mathbb{1}(d=j) | \cdot]$ can be expressed as functions of the choice probabilities. We denote these functions by $e(d; P[x])$. Let us substitute these functions into (4) and stack the equations for each possible value of x . The resulting system of equations can be solved for the social surpluses, and the $M \times 1$ solution vector S^π can be written compactly in matrix notation as follows:

$$S^\pi = \left(I_M - F^U(P) \right)^{-1} \mathbf{1} \otimes \mathbf{P}(d) \otimes [u(d) + e(d; P)], \quad (5)$$

where \otimes is the Hadamard (or element-by-element) product; \mathbf{P} is the $M(J-1) \times 1$ vector of conditional choice probabilities; $\mathbf{P}(d)$, $u(d)$ and $e(d; P)$ are $M \times 1$ vectors which stack the corresponding elements at all states for alternative d ; and $F^U(P)$ is the $M \times M$ matrix of unconditional transition probabilities induced by \mathbf{P} . If $F(d)$ is the $M \times M$ matrix of conditional transition probabilities $f(x^0 | x; d)$, then $F^U(P) = \sum_{d \in D} \mathbf{P}(d) \otimes F(d)$. Note that at each state there are only $J-1$ free choice probabilities. In our notation, probability vectors have dimension $J-1$. Without loss of generality, we exclude the probabilities of alternative 1 which are implicitly defined from the others.

2.2 Smooth policy iteration

We now introduce a Policy Iteration operator \mathcal{A} in the space of conditional choice probabilities. This operator plays a central role in our estimation algorithm. Using the same matrix notation of the previous section, let the $M(J-1) \times 1$ vector $\mathbb{1}$ denote a point in the space of conditional choice probabilities at all observable states; and let the $M \times 1$ vector S denote a point in the space of surplus values at all states. The Policy Iteration operator \mathcal{A} is defined as a composite mapping:

$$\mathcal{A}(\mathbb{1}) = \mathbf{1} \otimes \mathcal{A}(\mathbb{1}) \quad (6)$$

(iv) The Jacobian matrices of $\mathbb{3}_4(\cdot)$ and $\mathbb{a}(\cdot)$ are zero at the fixed point P .

Result (i) implies that the fixed point condition $P = \mathbb{a}(P)$ implicitly defines the optimal choice probabilities. By (ii), the iterations $\mathbb{1}_K = \mathbb{a}(\mathbb{1}_{K-1})$ define an algorithm which will always converge to the optimal choice probabilities. Result (iii) shows that this algorithm is equivalent to Newton's algorithm applied to the functional equation (2) which defines the surplus S^a . The first part of (iv) establishes that at the fixed point it is not possible to "improve" social surpluses by changing the conditional choice probabilities; that is, the optimal choice probabilities maximize the valuation operator locally. The second part of result (iv) plays an important part in establishing the properties of the estimation algorithm we propose in this paper.

We end this section with remarks on the role of Assumptions 1-3. Assumption 3 allowed us to use compact matrix notation, but it is not essential. On the contrary, Assumptions 1 and 2 are central to our analysis. Assumption 2 (Conditional Independence) is necessary for future utility differences not to depend on current unobservable state variables. Together with Assumption 1 (Additivity), this plays a crucial role in Hotz & Miller's representation of the value functions which we use to define the Policy Iteration operator. Assumption 1 can be relaxed to allow for multiplicative separability between observable and unobservable components. The differentiability in Assumption 2 is needed in order to make the Policy Iteration operator a differentiable function of choice probabilities.

3 Maximum likelihood estimation and nested algorithms

Let μ_u , μ_g and μ_f be the vectors of unknown parameters in the utility function u , the density of unobservables g and the conditional transition probability function f , respectively. That is, $\mu = (\mu_u; \mu_g; \mu_f)$. In order to guarantee the existence, consistency and asymptotic normality of the ML estimator, we impose smoothness of the primitives with respect to μ .

ASSUMPTION 4: $u(x; d; \mu_u)$, $g(x^0; x^0; \mu_g)$ and $f(x^0; x; d; \mu_f)$ are continuous and twice differentiable with respect to μ .

Suppose our data set consists of a cross-section of observations from a random sample of individuals $\{x_i; d_i; x_i^0 : i = 1, \dots, n\}$. Under Assumption 2, the log-likelihood function of the DDP model can be decomposed into conditional choice probability and transition probability terms as follows:

$$l(\mu) = l_1(\mu) + l_2(\mu_f) = \sum_{i=1}^n \ln P(d_i; x_i; \mu) + \sum_{i=1}^n \ln f(x_i^0; x_i; d_i; \mu_f) \quad (9)$$

Under Assumption 2, consistent estimates of the conditional transition probability parameters μ_f can be obtained from transition data without having to solve the Markov decision model. In the rest of the paper we focus on partial Maximum Likelihood estimation of $\mu = (\mu_u; \mu_g)$ based on the likelihood $l_1(\mu)$, given consistent estimates of μ_f obtained from likelihood $l_2(\mu_f)$.⁷ This two-stage estimation strategy, which reduces the computational burden of estimation, was used in Rust (1987) and Hotz and Miller (1993).⁸

Let μ^* denote the true value of μ hereafter. The (partial) MLE of μ^* can be computed using Rust's well known Nested Fixed Point Algorithm (NFXP). In this procedure, an 'inner' algorithm iterates on μ to compute the conditional choice probabilities $P(\mu) = P(d; \mu; P(\mu))$ and their derivatives for given parameter values.⁹ The 'outer' algorithm feeds on this solution and maximizes the likelihood using the BHHH method. We propose an alternative nested procedure:

Nested Pseudo Likelihood Algorithm (NPL).

Let $\hat{\mu}_f$ be an estimate of μ_f . Start with an initial guess for the conditional choice probabilities, $\mu^0 \in [0; 1]^{M \times J}$. At iteration $K = 1$, apply the following three steps:

⁷The discount factor β is assumed to be known.

⁸See Rust (1987, 1994) for a 3-stage extension which is asymptotically equivalent to full maximum likelihood.

⁹See Rust (1988) for a description of NFXP. The inner algorithm is a hybrid fixed point algorithm which begins with successive approximation iterations in the surplus function and then switches to Newton-Kantorovich iterations in order to guarantee convergence. Results (ii) and (iii) in our Proposition 1 suggest that the initial successive approximation iterations are actually not necessary for convergence.

Step 1: Given β^{K_i-1} , construct the smooth Policy Iteration mapping $a(\theta; \hat{\mu}_f; \beta^{K_i-1})$ described in section 2.2.

Step 2: Obtain a new pseudo-likelihood estimate of θ , θ^{K_i} , as

$$\theta^{K_i} = \arg \max_{f \in \mathcal{G}} \sum_{i=1}^n \ln a(d_i; x_i; \theta; \hat{\mu}_f; \beta^{K_i-1}) \quad (10)$$

Step 3: Update β using the 'arg max' from step 2, i.e.

$$\beta^{K_i} = a(\theta^{K_i}; \hat{\mu}_f; \beta^{K_i-1}); \quad (11)$$

Iterate in K until convergence in β (and θ) is reached.

In our nested procedure we swap the order of the two algorithms. That is, the outer algorithm iterates on a to solve the fixed point problem, and the inner algorithm maximizes a pseudo-likelihood function. In the inner algorithm (steps 1 and 2), a fixed β is used to construct the pseudo social surplus $\mathcal{S}(\theta; \hat{\mu}_f; \beta)$ and choice probabilities $a(\theta; \hat{\mu}_f; \beta)$ which approximate the exact choice probabilities $P(\theta; \hat{\mu}_f)$. The outer algorithm (step 3) updates the value of β and checks for convergence to a fixed point. However, note that each outer policy iteration is performed with updated parameter values. We will show that, as the outer policy iterations converge, the pseudo-social surplus converges to the true social surplus and the maximized pseudo-likelihood values converge to the true likelihood values.

Proposition 2 Equivalence of NFXP and NPL:

Suppose the pseudo-likelihood maximization problems in (10) have unique interior solutions for all K : Then, if the Nested Pseudo Likelihood algorithm (NPL) converges it does so to a root of the likelihood equation.

An implication of Proposition 1 is that $\theta^* P(\mu) = \theta^* \mu^0 = a(\mu; P(\mu)) = \theta^* \mu^0$; therefore, the pseudoscore is exactly equal to the score when the pseudoscore is evaluated at the fixed point of the policy iteration operator. But if the problem is sufficiently smooth, then on convergence NPL delivers a fixed point pair; that is, $\beta_{NPL} = a(\theta_{NPL}; \hat{\mu}_f; \beta_{NPL})$. It follows that the NPL algorithm has found a root of the likelihood equation, which is exactly what the NFXP algorithm does.

Notice that convergence of NFXP is guaranteed whereas convergence of NPL has not been proved. The result in Proposition 2 is conditional on convergence of NPL.¹⁰ Equations (10) and (11) which describe the NPL algorithm define an operator in the space of conditional choice probabilities: $\hat{A}(\cdot) = \hat{a}(\hat{\theta}(\cdot); \hat{\beta}_f; \cdot)$ where $\hat{\theta}(\cdot)$ is the maximizer of the pseudo-likelihood. Let $\hat{\theta}$ denote a root of the likelihood equations, and $\hat{\beta}$ the corresponding choice probabilities. It is clear from the proof of Proposition 2 that $\hat{\beta}$ is a fixed point of the $\hat{A}(\cdot)$ operator and $\hat{\theta} = \hat{\theta}(\hat{\beta})$. Therefore, convergence of NPL to a root of the likelihood equations amounts to convergence of fixed point iterations for $\hat{A}(\cdot) = \cdot$.¹¹ Theorem 5.4.2 of Judd (1998) gives sufficient conditions for local convergence of fixed point iterations. The critical condition is that the spectral radius of the Jacobian of \hat{A} at the fixed point should be less than 1, so that \hat{A} behave locally as a contraction. Since \hat{A} is a random operator, its spectral radius is a random variable and the condition for local convergence may not hold for every sample. However, we can show that the Jacobian of \hat{A} and its spectral radius converge in probability to zero. Therefore, local convergence is very likely in large enough samples. The next Proposition formalizes this result.

Proposition 3 Local convergence of NPL in large samples

Consider the NPL operator $\hat{A}_n(\cdot) = \hat{a}(\hat{\theta}_n(\cdot); \hat{\beta}_f; \cdot)$, where $\hat{\theta}_n(\cdot)$ is the maximizer of the pseudo log-likelihood defined in eq. (10) and is assumed to be interior and unique for all \cdot and for any sample of size n . Let $\hat{\theta}_n$ be a root of the likelihood equations, and $\hat{\beta}_n = \hat{\beta}(\hat{\theta}_n; \hat{\beta}_f)$. Let p_n be the probability that the NPL algorithm converge locally to $\hat{\beta}_n$, i.e., $\cdot^K = \hat{A}_n(\cdot^{K-1})$ converges to $\hat{\beta}_n$ in K if \cdot^0 sufficiently close to $\hat{\beta}_n$. Then p_n converges to 1 under the regularity conditions of Proposition 4.

We now give an example which illustrates the policy iteration operator and points at some likely determinants of the relative speed of the NPL and NFXP algorithms. Consider a class of models where: (1) The unobservable θ 's have independent across

¹⁰However, neither NFXP nor NPL guarantees convergence to a global maximum of the likelihood. In both cases, the researcher should start the algorithms from different points and keep track of the maxima.

¹¹We are implicitly assuming that convergence of NPL's inner algorithm is not an issue. NFXP converges if the BHHH method with optimal step is used in its outer algorithm. The same method can be used in NPL's inner algorithm.

alternatives extreme value distributions; and (2) there is multiplicative separability between x and θ in the utilities, i.e., $u(d; \theta) = H(d)\pm(\theta)$ where $H(d)$ is an $M \times J$ matrix with rows $h(x; d)^0$ and $h(\cdot)$, $\pm(\cdot)$ are known vector-valued functions. In this case, $\pi(\cdot)$ is a well known (logistic) function of choice-specific utilities and $e(d; \cdot) = c_j \ln \pi_j(d)$ where c stacks Euler's constant. The probabilities that enter the pseudo-likelihood function are

$$p_j(d; x; \theta; \mu_f; \cdot) = \frac{\exp\{y_0(x; d; \mu_f; \cdot) + y_1(x; d; \mu_f; \cdot)\pm(\theta)\}g_j}{\sum_{j=1}^J \exp\{y_0(x; j; \mu_f; \cdot) + y_1(x; j; \mu_f; \cdot)\pm(\theta)\}g_j}$$

where $y_0(x; d; \mu_f; \cdot)$ is an element of the vector $-F(d) \mathbf{1}_M \mathbf{1}_J^{-1} F^U(\cdot) \mathbf{1}_J^{-1} \mathbf{P}_j(\cdot)$ and $y_1(x; d; \mu_f; \cdot)$ is a row of the matrix $H(d) + F(d) \mathbf{1}_M \mathbf{1}_J^{-1} F^U(\cdot) \mathbf{1}_J^{-1} \mathbf{P}_j(\cdot)$

This model has several features that make the use of NPL particularly attractive. First, since the vectors $y(\cdot)$ do not depend on θ they are fixed over a pseudo-likelihood estimation. To obtain $p_j(d; x; \theta; \mu_f; \cdot)$ for different values of θ we do not have to repeat the inversion and multiplication of large matrices which is required for policy valuation. Second, the extreme value assumption implies that integration over unobservables during a pseudo-likelihood estimation has a simple closed form. And third, the pseudo-likelihood function is globally concave in $\pm(\theta)$, which guarantees convergence of the hill-climbing pseudo-likelihood iterations for any initial value of θ .¹² In contrast, to compute the probabilities $P(d; x; \theta; \mu_f)$ that enter the likelihood function we have to invert and multiply $M \times M$ matrices repeatedly. Furthermore, $P(d; x; \theta; \mu_f)$ is not globally concave in θ , nor in a transformation of θ . Therefore, convergence of NFXP's outer BHHH algorithm may require the use of optimal steps with a significant increase of the computational cost of estimation.¹³ In Section 5 we show that NPL is indeed much faster than NFXP for Rust's bus replacement problem, which is an example of this class of models.

¹²Given the ML estimator of $\pm(\theta)$, we can obtain an estimator of θ using minimum distance. This is the ML estimator of θ if $\pm(\cdot)$ is a one-to-one function.

¹³Without multiplicative separability in the utility function, the pseudo likelihood function will not in general be globally concave. Furthermore, to obtain the choice-specific utilities $v(x; d; \theta; \mu_f; \cdot)$ for different values of θ we have to multiply $M \times M$ matrices. However, it is still true that we do not have to invert the matrix $\mathbf{1}_J^{-1} F^U(\cdot)$, which is the main computational cost of policy valuation.

In general we think that the main factor affecting the relative speed of the two algorithms is the cost of integration over unobservables in $\pi(\cdot)$ and $e(\cdot)$, relative to the cost of inverting large matrices for policy valuation. NFXP and NPL require integration over unobservables and matrix inversion at each policy iteration. But additional integration over unobservables is also required at pseudo-likelihood climbing iterations. To the extent that it saves on policy iterations, NPL will reduce the number of matrix inversions, but not necessarily the number of integrations over unobservables. In models where integration over unobservables is the main component of computational cost NPL may not reduce the cost of estimation.

Finally, note that result (iv) in Proposition 1 is crucial to obtain equivalence of NFXP and NPL. Based on this, it is straightforward to see that the equivalence result also holds for full likelihood versions of NFXP and NPL.¹⁴ Also note that finite horizon models are covered in our infinite horizon framework if the decision period t is included among the observable state variables.¹⁵ Furthermore, the NPL algorithm defines a recursive extension of Hotz and Miller's Conditional Choice Probability (CCP) estimator. We turn to this issue in the next section.

4 Sequential policy iteration estimators

Let $\hat{\beta}_F$ denote a consistent estimator of conditional transition probability parameters, and let \hat{P}_i^0 be a consistent, nonparametric estimator of the true conditional choice probabilities P^π . Consider using \hat{P}_i^0 as an initial guess in our NPL algorithm. Performing one, two and in general K iterations of the NPL algorithm yields a sequence $\hat{\theta}^1; \hat{\theta}^2; \dots; \hat{\theta}^K$ of statistics which can be used as estimators of θ^π . We call them sequential Policy Iteration (PI) estimators. Thus, for $K \geq 1$ the K th stage PI

¹⁴See Appendix 1. However, our discussion above suggests that the potential gains from using NPL instead of NFXP are much larger in a partial likelihood context.

¹⁵The inverse of the matrix $(I - F^U)$ which is used in the valuation operator has a very special structure. Therefore, it is harder to conjecture what the potential gains of using NPL are in this case. However, all the results in Propositions 1 - 4 of this paper apply, beginning with the key 'zero-jacobian' property of Proposition 1 (iv).

estimator is defined as:

$$\hat{\theta}^K = \arg \max_{\theta \in \Theta} \sum_{i=1}^n \ln \pi^a(d_i; x_i; \theta; \hat{p}_f; \hat{\mu}_f^{K-1}) \quad (12)$$

where, for $K \geq 1$, $\hat{\mu}_f^K = \pi^a(\hat{\theta}^K; \hat{p}_f; \hat{\mu}_f^{K-1})$, and $\hat{\mu}_f^0$ is a nonparametric estimator of P^a .

It is clear from Proposition 3 that as the NPL algorithm converges in $\hat{\mu}_f^K$ the corresponding K -stage PI estimators converge to the partial MLE. The family of PI estimators encompasses also the Conditional Choice Probabilities (CCP) estimator proposed by Hotz and Miller (1993). The CCP estimator is defined as the value of θ that solves the system of equations:

$$\sum_{i=1}^n \sum_{j=1}^J Z_i^j (I[d_i = j] - \pi^a(j; x_i; \theta; \hat{p}_f; \hat{\mu}_f^0)) = 0 \quad (13)$$

where fZ_i^j are vectors of instrumental variables (e.g., functions of x_i). It is straightforward to verify that the 1 -stage PI estimator is a CCP estimator where $Z_i^j = I[d_i = j] - \pi^a(j; x_i; \theta; \hat{p}_f; \hat{\mu}_f^0)$.¹⁶

In this section we study the asymptotic statistical properties of this sequence of estimators. The main result is in Proposition 4 which shows that for any value of K the PI estimators are consistent and asymptotically equivalent to the partial MLE of θ^a . More formally:

Proposition 4:

Let P^a be the true probability distribution of d conditional on x , and let f^a be the true conditional transition probability of x . Define $\theta^a = (\mu_f; p_f)$, and let $\Theta \subset \mathbb{R}^{\mu_f} \times \mathbb{R}^{p_f} \subset [0; 1]^{(J+1)M}$ be the set of possible values of θ . Consider the following regularity conditions.

¹⁶As seen in section 2, we use Hotz and Miller's representation of the value function to construct choice probabilities. Our policy iteration operator π^a adapts their representation to our 1-horizon framework and makes explicit its dependence on guesses of the choice probabilities. Also, note that in their paper Hotz and Miller proposed and used yet another representation of the value function for problems with terminal histories. This representation is also based on the invertibility of the mapping from utility differences to choice probabilities. However, for arbitrary guesses of the choice probabilities it is different from the first one. Our Propositions 2 and 3 do not apply to CCP estimators based on this second representation.

- (i) E_{θ} and E_f are compact sets.
- (ii) $a(d; x; \theta)$ is continuous and twice continuously differentiable in θ .
- (iii) $a(d; x; \theta) > 0$ for any $(d; x) \in D \times X$ and $\theta \in \Theta$.
- (iv) $f(d_i; x_i^0; x_i^1)$ for $i = 1, 2, \dots, n$ are independently and identically distributed, and $\Pr(x_i = x^m) > 0$ for any $x^m \in X$.
- (v) There is a unique $\mu_f^* \in E_f$ such that $f(x^0; x; d; \mu_f^*) = f^*(x^0; x; d)$ for all $(x^0; x; d)$.
- (vi) There is a unique $\theta^* \in E_{\theta}$ such that, for any $(d; x) \in D \times X$, $P(d; x; \theta^*; \mu_f^*) = P^*(d; x)$. Furthermore, for any $\theta \in E_{\theta}$ the set $\{(d; x) : a(d; x; \theta; \mu_f^*) \in P^*(d; x)\}$ is non empty.
- (vii) $(\theta^*; \mu_f^*) \in \text{int}(E_{\theta} \times E_f)$
- (viii) $\hat{\mu}_f$ and $\hat{\theta}_i^0$ are consistent estimators of μ_f^* and P^* , respectively, and

$$\sqrt{n} \begin{pmatrix} \frac{1}{n} \sum_i \frac{\partial \ln a(d_i; x_i; \theta^*; \mu_f^*; P^*)}{\partial \theta} \\ \hat{\mu}_f - \mu_f^* \\ \hat{\theta}_i^0 - \theta^* \end{pmatrix} \xrightarrow{D} N(0; -)$$

Under these conditions, a policy iteration estimator, $\hat{\theta}^K$ for $K \geq 1$, is root-n-consistent, asymptotically normal, and asymptotically equivalent to the partial maximum likelihood estimator, i.e., $\sqrt{n}(\hat{\theta}^K - \theta^*) \xrightarrow{D} N(0; V^*)$ where

$$V^* = -j_{\theta\theta}^1 + -j_{\theta\theta}^1 H_f - f_{\theta} + -o_f H_f^0 + H_f V(\hat{\mu}_f) H_f^0 - j_{\theta\theta}^1 \quad (14)$$

and:

$$\begin{aligned} -j_{\theta\theta}^1 &= E \frac{\partial^2 \ln P_i^*}{\partial \theta^2} \frac{\partial \ln P_i^*}{\partial \theta} ; & H_f &= E \frac{\partial \ln P_i^*}{\partial \theta} \frac{\partial \ln P_i^*}{\partial \mu_f^*} \\ -f_{\theta} &= -j_{\theta f}^0 = E \frac{\partial \ln f_i^*}{\partial \mu_f^*} \frac{\partial \ln f_i^*}{\partial \theta} ; & & E \frac{\partial \ln f_i^*}{\partial \mu_f^*} \frac{\partial \ln P_i^*}{\partial \theta} \end{aligned}$$

with $P_i^* \sim P(d_i; x_i; \theta^*; \mu_f^*)$ and $f_i^* \sim f(x_i^0; x_i^1; d_i; \mu_f^*)$:

The K stage PI estimator is a particular case of a Quasi-Generalized M-estimator as defined in Gourieroux and Monfort (1995), with $(\mu_f^*; P^*)$ as the vector of nuisance parameters. The proof of Proposition 4 uses an induction argument. We show that if $\hat{\theta}^{K-1}$ is root-n-consistent and asymptotically normal then $\sqrt{n}(\hat{\theta}^K - \theta^*) \xrightarrow{D} N(0; V^*)$, and $\hat{\mu}_f^K \xrightarrow{P} (\theta^K; \hat{\mu}_f; \hat{\theta}^{K-1})$ is also root-n-consistent and

asymptotically normal. Since $\hat{\mu}_i^0$ is root-n-consistent and asymptotically normal by assumption, the proof is complete.¹⁷

The regularity conditions (ii) and (iii) follow from our assumptions in Sections 2 and 3. The assumption of iid observations in condition (iv) makes the proof of Proposition 4 easier, but it can be relaxed.¹⁸ Conditions (v) and (vi) are identification assumptions. In particular, (vi) implies that μ_i^{α} uniquely maximizes in μ both $E \ln^{\alpha}(d_i; x_i; \mu; \mu_f^{\alpha}; P^{\alpha})$ and $E \ln P(d_i; x_i; \mu; \mu_f^{\alpha})$. That is, condition (vi) implies identification of μ^{α} in the context of partial maximum likelihood estimation and in the context of the pseudo-likelihood estimations needed to obtain the K_i stage estimators. Without further assumptions on $\mu(\cdot)$, identification in maximum likelihood does not imply identification in the pseudo-likelihood.

Notice that the asymptotic variance of the K_i stage estimators does not depend on the variance of the initial nonparametric estimator of the choice probabilities, $\hat{\mu}_i^0$. This asymptotic property of the PI estimators results from Proposition 1. In particular, Proposition 1 establishes that for any pair $(d; x)$, $\frac{\partial}{\partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0} = 0$, and using the information matrix equivalence this implies that:

$$E \frac{\frac{\partial^2}{\partial \mu \partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0}}{\frac{\partial^2}{\partial \mu \partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0}} = E \frac{\frac{\partial}{\partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0} \frac{\partial}{\partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0}}{\frac{\partial^2}{\partial \mu \partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0}} = 0$$

In a Taylor's expansion of the first order conditions that define a K_i stage PI estimator, the term associated with $\hat{\mu}_i^{K_i-1}$ is

$$\frac{\partial}{\partial \mu} \ln^{\alpha}(d_i; x_i; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0} \frac{\partial}{\partial \mu} \ln^{\alpha}(d_i; x_i; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0} P_{n,i}^{-1}(\hat{\mu}_i^{K_i-1}; P^{\alpha})$$

Therefore, if $(1/n) \sum_{i=1}^n \frac{\partial^2}{\partial \mu \partial \mu} \ln^{\alpha}(d_i; x_i; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0}$ converges in probability to $E \frac{\partial^2}{\partial \mu \partial \mu} \ln^{\alpha}(d; x; \mu; \mu_f^{\alpha}; P^{\alpha})|_{\mu=\mu_i^0}$ and $P_{n,i}^{-1}(\hat{\mu}_i^{K_i-1}; P^{\alpha})$ converges to a vector of

¹⁷Given our assumptions that x is discrete and $\Pr(x_i = x^m) > 0$ for any $x^m \in X$, obtaining P_n^{-1} -consistent non parametric estimators is straightforward. Delgado and Mora (1995a,b) show that in non-parametric regression with discrete regressors frequency or 'cell' estimators and nearest-neighbor estimators are P_n^{-1} -consistent under very weak conditions. Frequency estimators for empty x -cells in a given sample are defined to be zero. Bierens (1983) proves that in a non-parametric regression with discrete explanatory variables, the kernel estimator is root-n-consistent.

¹⁸We consider a cross-section for the sake of simplicity, but our results here and in Section 3 of the paper can be extended to the case of panel datasets as long as the conditional independence assumption holds along both dimensions of the panel.

random variables with finite variances, then this term does not have any effect on the asymptotic distribution of $\hat{P}_n^{(K; \mu^a)}$.

Given the asymptotic equivalence of the PI estimators and the partial MLE, the analysis of the relative performance of these estimators in finite samples is of interest. Unless there is a loss of precision, why not use the computationally inexpensive 1-stage PI estimator? An important limitation of Hotz-Miller's estimator is that initial nonparametric estimates of the conditional choice probabilities can be very imprecise, and this lack of accuracy is transmitted to the estimates of the structural parameters (see Eckstein and Wolpin, 1989, Rust, 1994, and Hotz et al., 1994). Our K -stage estimator may overcome this problem by iterating K times in the smooth policy iteration operator. The computational cost of these $(K - 1)$ additional iterations is equal to the cost of $(K - 1)$ policy iterations and $(K - 1)$ pseudo-maximum likelihood estimations. Therefore, for intermediate values of K we get asymptotically equivalent estimators which are still cheaper to obtain than the MLE yet potentially more precise in finite samples than the Hotz-Miller estimator. Is a 2-stage or 3-stage estimator enough to obtain significant gains in precision with respect to Hotz-Miller? We address this issue with a Monte Carlo experiment in Section 5.2.

Using consistent estimators of μ^a and μ_f^a it is straightforward to obtain a consistent estimator of V^a . Furthermore, though V^a depends on expectations that involve partial derivatives for $P(\mu^a)$, by Proposition 1 it is possible to estimate consistently V^a by using $\frac{\partial}{\partial \mu^a}(d_i; x_i; \mu^K; \hat{\mu}_f; \hat{\mu}_f^{K-1}) = \frac{\partial}{\partial \mu^a}$ instead of $\frac{\partial}{\partial \mu^a} P(d_i; x_i; \mu^K; \hat{\mu}_f) = \frac{\partial}{\partial \mu^a}$. That is, the estimation of V^a does not require one to solve once the fixed point problem to obtain the vector of choice probabilities $P(\mu^K; \hat{\mu}_f)$ and its partial derivatives.

It is simple to verify that Proposition 4 can be extended to a full maximum likelihood context, i.e., the joint estimation of μ^a and μ_f^a . However, as mentioned in Section 3 we focus on partial maximum likelihood estimation because we believe it is in this context that the potential computational gains of the NPL algorithm and PI estimators are greatest.

5 The performance of the NPL algorithm and PI estimators

5.1 Relative speed of NPL and NFXP

In order to illustrate the performance of our NPL algorithm in maximum likelihood estimation, we use Rust's well known bus replacement model and dataset.¹⁹ We obtain partial ML estimates using the NPL and the NFXP algorithms for different specifications of the model according to the dimension of the state space (i.e., from 100 to 1100 cells) and the number of structural parameters (i.e., 2 and 4). For a detailed technical description of the two algorithms used in this section see Appendix 2.²⁰

Clearly, the CPU time required by these algorithms will depend on our choice of the values that initialize them, i.e., initial guesses of structural parameters for NFXP and initial conditional choice probabilities for NPL. In order to make the initial values for the two algorithms comparable, we consider a researcher who obtains these guesses from the data using a Hotz-Miller estimator. That is, the initial vector of structural parameters for NFXP is the Hotz-Miller estimator $\hat{\beta}_1$, and the initial conditional choice probabilities for NPL are the ones obtained from the Hotz-Miller estimator $\hat{\beta}_1$ ($\hat{\beta}_1; \hat{\pi}_0$).

We summarize our results for the 2 and 4-parameter specifications in Figures 1 and 2, respectively. Figures 1A and 2A plot the fraction of total estimation time which is spent in policy iterations against the number of points in the state space. We see that even when the number of cells in the state space is relatively small policy iterations represent almost 100% of CPU time. This is the case for both algorithms, though the ratio tends to increase more slowly for the NPL algorithm.²¹ An important feature of

¹⁹See Rust (1987). Rust's model has been used in other studies to evaluate the performance of alternative algorithms and estimators, e.g. Hotz et al. (1994) and Rust (1997b).

²⁰It is important to underline that the NFXP algorithm that we use has two features which contribute very significantly to improve its computational efficiency (both features have been considered by Rust, 1987, 1988). First, we use a closed-form expression for the gradient of the likelihood function. Second, at each outer iteration, we use "smart guesses" for the vector of choice probabilities that initialize the policy iteration algorithm. For further details see Appendix 2.

²¹The reason is that the number of policy iterations under NPL is smaller than under NFXP. See

this example is that, in the context of partial likelihood estimation, policy iterations are much more expensive than pseudo-likelihood climbing iterations. Notice that both types of iterations use the policy iteration operator $T^a(\mu; \psi)$. However, pseudo-likelihood climbing iterations with fixed ψ at each stage do not require repeated matrix inversions in order to compute the policy valuation operator. Therefore, it is quite relevant to assess to what extent the use of NPL instead of NFXP reduces the number of policy iterations in the estimation procedure. This is shown in Figures 1B and 2B. Notice that, in our example, the size of the state space does not affect the number of policy iterations in any of the two algorithms. For the model with two parameters the ratio of the number of policy iterations of NFXP and NPL is 5:5, and for the model with four parameters this ratio is equal to 9. Therefore, we may already conclude that in these examples the NPL is 5:5 and 9 times faster, respectively, than the NFXP algorithm. Figures 1C and 2C plot the logarithm of CPU time against the number of cells in the state space.

The use of the Hotz-Miller estimator to produce 'comparable' initial guesses actually undervalues the relative performance of the NPL algorithm. In our example, the Hotz-Miller estimator produces accurate initial guesses. When we considered cases in which initial guesses are poor, our results were more favorable to the NPL algorithm. Even when we used extremely biased and imprecise initial values for the choice probabilities (e.g., random draws from a $U(0; 1)$), the NPL algorithm always converged to the MLE. In contrast, the NFXP algorithm did not converge for a wide range of initial values of the structural parameters.²² And when it converged, the number of policy iterations increased relative to NPL with equally arbitrary initial guesses. For instance, when we used a vector of zeros as an initial guess for the NFXP algorithm and random draws from a $U(0; 1)$ to initialize NPL, the ratio of policy iterations NFXP/NPL was always between 10 and 15 depending on the number of parameters

Figures 1B and 2B.

²²That was the case even for some values within a 1% confidence region around the ML estimate. NFXP will always converge if we compute an optimal step in the hill climbing iteration. However, the use of an optimal step requires one to solve the dynamic programming problem several times at each hill climbing iteration, and this increases enormously the computational cost of this algorithm.

and the size of the state space.

A final warning seems appropriate given the limited scope of this experiment. Our example consists of an infinite horizon, binary choice model with extreme value distributed unobservables and multiplicative separability between parameters and state variables in the utility functions. The relative speed of NPL and NFXP could be different for other specifications.

5.2 The precision of PI estimators: A Monte Carlo exercise

In section 4 we established that all K-stage PI are asymptotically equivalent. We now look into their behavior in finite samples. The Hotz-Miller estimator is known to be quite sensitive to the quality of the initial non-parametric estimates of choice probabilities. This suggests that there may be a trade-off between increasing computational cost and increasing precision as we move along the sequence of estimators. In order to illustrate and analyze this trade-off in finite samples, we carried out the following Monte Carlo experiment. We used Rust's bus engine model with parameters equal to ML estimates as the DGP. For sample sizes 1000, 5000 and 10000 we generated 1000 samples and for each of them we obtained the sequence of PI estimators and its limit, the ML estimator.²³ The size of the state space grid is 200. Recall from Figures 1B and 2B that the choice of grid size hardly affects the number of policy iterations needed to obtain the MLE.

Tables 1 to 3 present summary statistics for the three experiments. For the 1, 2 and 3 -stage PI estimators, we report in these tables the mean and median of the absolute estimation error and the empirical standard deviations, in all cases relative to the corresponding statistic for the MLE. In order to take into account the potential sensitivity of the results to outliers we also report results for trimmed empirical distributions of the estimators, i.e., estimates between percentiles 5 and 95.

²³Each sample has been obtained in the following way. First, we obtain the steady-state distribution of the observable state variable (cumulative mileage) using transition probabilities and true choice probabilities. Second, from this distribution we obtain n random draws of the state variable, where n is the sample size. Finally, using the optimal decision rule and simulations of the unobservable epsilons, we obtain the optimal decisions associated with each of the n simulations of the state variables.

In the three experiments, we find very significant benefits of doing more than one policy iteration, in particular for relatively large sample sizes. In general, the 1-PI estimator performs poorly relative to the other estimators. More interestingly, the relative performance of the 2-PI is excellent. Even for very small sample sizes (and, consequently, imprecise kernel estimates of the choice probabilities) the 2-PI estimator is very similar to MLE. Most of the benefits of additional policy iterations are obtained when one goes from 1 to 2 iterations. In Table 4 we consider the performance of the three estimators when the initial values for the choice probabilities are very imprecise. Instead of kernel estimates we use as initial probabilities the kernel estimates plus a noise term.²⁴ This results in artificially poor initial values for the choice probabilities. Figure 3 plots the true choice probability, the kernel estimates and the noisy estimates for one of the replications. The results of this experiment actually reinforce the ones from Experiments 1 to 3. Poor initial guesses have a very serious effect on the 1-PI estimator but a relatively small effect on the 2-PI, which is still close to the MLE.

Therefore, when we construct expected value functions using choice probabilities which are closer to the fixed point our estimates become much more precise. In this application, two policy iterations are enough to get choice probabilities which are close enough to the fixed point. A possible interpretation of this finding is based on the qualitative differences between going from 1 to 2 policy iterations and going from 2 to more iterations. The second policy iteration is the first one in which we incorporate the structure of the model to obtain the choice probabilities used to compute the pseudo-social surplus. These probabilities incorporate for the first time the following information which is not contained in the initial nonparametric estimates: (1) a distributional assumption about the unobservable state variables; (2) a parametric

²⁴The noisy estimator for choice probability $P(x)$ is:

$$P(x) = \hat{P}(x) + \frac{k}{n^2} e(x)$$

where $\hat{P}(x)$ is the kernel estimator; k is an arbitrary constant; n is the sample size; and $e(x)$ is a random variable uniformly distributed in the interval $[\min\{\hat{P}(x), 1 - \hat{P}(x)\}, \max\{\hat{P}(x), 1 - \hat{P}(x)\}]$, where $\min\{\hat{P}(x), 1 - \hat{P}(x)\}$ is equal to the minimum of $\hat{P}(x)$ and $1 - \hat{P}(x)$. Notice that this estimator has the same asymptotic properties as the kernel estimator.

assumption for the one-period utility function; (3) the assumption of additivity between observables and unobservables in the one-period utility function; and (4) the assumption of additivity between current and future utilities in the intertemporal utility function. All subsequent policy iterations impose (recursively) that (5) the choice probabilities should be a fixed point of the policy iteration mapping, but no further assumptions about the functional form of the primitives are incorporated.

In Table 5 we study the discrepancy between the empirical standard deviations of the estimators in the Monte Carlo distributions and the average estimate of their asymptotic standard error. We can see that for the 2-stage and 3-stage this discrepancy is small and of the same order of magnitude as for the MLE. However, for the 1-stage estimator the estimated asymptotic standard error is clearly downward biased.

6 Concluding remarks

We have proposed a new algorithm to obtain Maximum Likelihood estimates of dynamic programming discrete choice models. We have shown that this algorithm may produce the MLE using significantly fewer policy iterations than the usual Nested Fixed Point Algorithm and that this results in large computational gains in an example taken from a special class of infinite-horizon, partial likelihood problems. Based on our algorithm, we have defined a family of sequential Policy Iteration estimators which encompass Hotz-Miller's CCP estimator and the ML estimator as extreme cases. All K -stage PI estimators are consistent and asymptotically equivalent to the partial ML estimator. Our Monte Carlo experiments confirm the relatively poor performance of the 1-stage PI estimator when the nonparametric estimates of conditional choice probabilities are imprecise. However, the performance of the 2-stage PI estimator in all our experiments is excellent. Its finite sample variance and bias are very similar to those of the MLE. Since this estimator is computationally much cheaper to obtain than MLE, our ability to obtain precise estimates of interesting economic models with more than two or three state variables might be significantly enhanced. We would like to see these encouraging results confirmed by the use of

2-stage and k-stage PI estimators in actual applications.

APPENDIX 1. Proof of Propositions

Proofs of Proposition 1:

Proof of (i): P is the unique fixed point of the policy iteration operator \mathcal{A} .

We first establish that the surplus function $S^*(\cdot)$ which solves the functional equation (2) is unique. Let $\mathcal{C}(\cdot)$ be the smooth Bellman operator defined on the right hand side of equation (2). Wozniakowski, Traub and Rust (1998) show (pg. 9) that the smooth Bellman operator belongs to the class of quasilinear mappings (Definition 1), and that all quasilinear mappings are in fact contraction mappings (Theorem 1). Uniqueness of $S^*(\cdot)$ follows from the contraction mapping theorem.

Second, we show that if ψ^* is a fixed point of $\mathcal{A}(\cdot)$, then $\mathcal{H}(\psi^*)$ is a fixed point of $\mathcal{C}(\cdot)$. Notice that by the definition of the mapping $\mathcal{C}(\cdot)$ in equation (2) we have that:

$$\mathcal{C}(S) = \int_{d \in D} \alpha(S) \alpha [u(d) + e(d; \alpha(S)) + \beta F(d)S]$$

Let ψ^* be a fixed point of $\mathcal{A}(\cdot)$, i.e., $\psi^* = \alpha(\mathcal{H}(\psi^*))$. Then,

$$\begin{aligned} \mathcal{C}(\mathcal{H}(\psi^*)) &= \int_{d \in D} \alpha(\mathcal{H}(\psi^*)) \alpha [u(d) + e(d; \alpha(\mathcal{H}(\psi^*))) + \beta F(d)\mathcal{H}(\psi^*)] \\ &= \int_{d \in D} \psi^*(d) \alpha [u(d) + e(d; \psi^*) + \beta F(d)\mathcal{H}(\psi^*)] = \mathcal{H}(\psi^*) \end{aligned}$$

Now suppose that ψ_1^* and ψ_2^* are different fixed points of $\mathcal{A}(\cdot)$. Since $\mathcal{H}(\psi_1^*)$ and $\mathcal{H}(\psi_2^*)$ are fixed points of $\mathcal{C}(\cdot)$, uniqueness of the fixed point of $\mathcal{C}(\cdot)$ implies that $\mathcal{H}(\psi_1^*) = \mathcal{H}(\psi_2^*)$. But then, $\psi_1^* = \mathcal{A}(\psi_1^*) = \alpha(\mathcal{H}(\psi_1^*)) = \alpha(\mathcal{H}(\psi_2^*)) = \mathcal{A}(\psi_2^*) = \psi_2^*$, a contradiction. Therefore, uniqueness of the fixed point of $\mathcal{A}(\cdot)$ follows from uniqueness of the fixed point of $\mathcal{C}(\cdot)$.

Finally, note that $P \in \alpha(S^*)$ and $S^* = \mathcal{H}(P)$. Therefore, $P = \alpha(\mathcal{H}(P)) = \mathcal{A}(P)$; i.e., $\mathcal{A}(\cdot)$ does have a fixed point, which is P .

To prove (ii) to (iv) we use three properties which are derived from the Williams-Daly-Zachary theorem and Proposition 1 in Hotz and Miller. Recall McFadden's social surplus function, which computes the surplus as a function of the vector of deterministic choice-specific utilities: $G(u) \in G(fu(j); j = 1; \dots; J) \in \mathbb{R} \max_j fu(j) + \sum_j g_j g(d^j)$.

Note that $S^x(x) = G(fv(x; j); j = 1; \dots; J; g)$. By the Williams-Daly-Zachary theorem: (1) $G()$ is globally convex; and (2) the gradient of $G()$ is equal to the vector of conditional choice probabilities (see Rust (1994), Theorem 3.1). Let $Q(\cdot)$ be the vector-valued function mapping $J - 1$ utility differences into $J - 1$ choice probabilities. By Proposition 1 in Hotz and Miller this mapping is invertible.

Property (a): Let $\frac{1}{4}$ be a $(J - 1) \times 1$ vector of choice probabilities, and $e(j; \frac{1}{4})$ the conditional expectation function defined in Section 2. Then,

$$\frac{\partial}{\partial \frac{1}{4}} \sum_{j=1}^J \frac{1}{4}_j e(j; \frac{1}{4}) = Q^{-1}(\frac{1}{4})$$

Proof: Consider the following representation of the surplus:

$$G(u) = \sum_j P(j) fu(j) + E["(j)j \pm (" = j]g = \sum_j Q_j(\Phi u) fu(j) + W_j(\Phi u)g$$

where Φu is the vector of utility differences $[u(2) - u(1); \dots; u(J) - u(1)]^0$, $Q_1(\Phi u) = 1 - \sum_{j=2}^J Q_j(\Phi u)$ and $W(\cdot)$ is the vector-valued function mapping $J - 1$ utility differences into J conditional expectations of unobservables. In matrix notation, we can write $G(u) = [1 - i^0 Q(\Phi u); Q(\Phi u)^0] (u + W(\Phi u))$. Differentiating with respect to $[u(2); \dots; u(J)]^0$ and applying the WDW theorem we get

$$[i - i; I] (u + W(\Phi u)) = i \frac{\partial Q(\Phi u)^0 \# i^{-1}}{\partial \Phi u} \frac{\partial W(\Phi u)^0 \bar{A}}{\partial \Phi u} [1 - i^0 Q(\Phi u)]^i \quad (\text{Ap.1})$$

We now differentiate $\sum_{j=1}^J \frac{1}{4}_j e(j; \frac{1}{4})$: Since the $Q(\cdot)$ mapping is invertible we can write $e(j; \frac{1}{4}) = W_j(Q^{-1}(\frac{1}{4}))$. Therefore,

$$\begin{aligned} \frac{\partial}{\partial \frac{1}{4}} \sum_{j=1}^J \frac{1}{4}_j e(j; \frac{1}{4}) &= \frac{\partial}{\partial \frac{1}{4}} \sum_{j=1}^J \frac{1}{4}_j W_j(Q^{-1}(\frac{1}{4})) = \quad (\text{Ap.2}) \\ &= [i - i; I] W(Q^{-1}(\frac{1}{4})) + \frac{\partial Q(Q^{-1}(\frac{1}{4}))^0 \# i^{-1}}{\partial \Phi u} \frac{\partial W(Q^{-1}(\frac{1}{4}))^0 \bar{A}}{\partial \Phi u} [1 - i^0 \frac{1}{4}]^i \end{aligned}$$

We now solve for $[i - i; I] W(\Phi u)$ in (Ap. 1) and substitute in (Ap. 2) to complete the proof.

Property (b): For any vector of utility differences Φu , the Jacobian matrix $\frac{\partial Q(\Phi u)}{\partial \Phi u}$ is positive semidefinite.

Proof: Notice that $\frac{\partial G(u)}{\partial u} = (1 - \beta Q(\Phi u) - Q(\Phi u))$, and $\frac{\partial^2 G(u)}{\partial u \partial u} = -\beta Q'(\Phi u) - Q'(\Phi u)$. Since the surplus function is convex it has a positive semidefinite Hessian. Therefore, $\frac{\partial Q(u)}{\partial \Phi u}$ is also positive semidefinite.

Property (c): Let λ be a $M(J+1) \times 1$ vector of choice probabilities at all states. λ_m is the column vector containing the $(J+1)$ choice probabilities in λ associated with state $x = x^m$. Consider the function $H(\lambda; S) = \sum_j \lambda_j [u(j) + e(j; \lambda) + \beta F(j)S]$ defined for arbitrary vectors of choice probabilities and surplus values at all states. Finally, let Φv_m be the vector of utility differences constructed from S :

$\Phi v_m(j) = (u(x_m; j) - u(x_m; 1)) + \beta (F(x_m; j) - F(x_m; 1))S$. Then,

- (c:1) $\frac{\partial H_n}{\partial \lambda_m} = 0$ for $n \neq m$
- (c:2) $\frac{\partial H_m}{\partial \lambda_m} = -\beta Q^{-1}(\lambda_m) + \Phi v_m$
- (c:3) H_m is a concave function of λ_m

Proof: Given that H_m does not depend on the probabilities for states different to m , (c.1) follows trivially. (c.2) results directly from Property (a), and (c.3) from Properties (b) and (c.2).

Proof of (ii): The sequence $\lambda_K = \alpha(\lambda_{K-1})$; $K = 1; \dots; 1$ converges to λ^* for any λ_0 .

Consider the Markov decision process defined as follows. (a) The state variable is the observable state vector of the original problem, x . (b) For any value of the state variables the action space is the interior of the J -dimensional simplex. (c) The current period return is $U(x; \lambda) = \sum_j \lambda_j [u(x; j) + e(j; \lambda)]$. (d) The transition probabilities $p(x^0 j x; \lambda) = \sum_j \lambda_j f(x^0 j x; j)$. We will show that in this transformed problem λ , α and α are indeed policy valuation, policy improvement and policy iteration operators. Checking that λ is the valuation operator for the transformed problem is trivial. Now notice that the policy improvement step involves solving $\max_{\lambda} H(\lambda; \lambda)$. By property (c.2), it is straightforward to see that $\alpha(\lambda)$ satisfies the first order conditions for an interior solution to this problem. Since the objective function is globally concave by Property (c.3), a solution does exist and it

is $\alpha(\cdot)$. Finally, notice that the transformed problem is a Markov decision process with discrete state space and continuous action space and that a maximizing decision rule always exists in the Bellman equation. Therefore, Theorem 6.4.6 in Puterman (1994) applies and policy iteration is guaranteed to converge.

Proof of (iii): Equivalence of α and Newton iterations

The surplus vector S^α solves the functional equation (4), $S = \mathcal{C}(S)$. Therefore, it is a zero of the operator $T(S) = S - \mathcal{C}(S)$; i.e., $T(S^\alpha) = 0$. Newton iterations are defined by $S^{K+1} = S^K + \mathcal{H}^{-1} T(S^K)$. Notice we can write the Bellman operator as $\mathcal{C}(S) = H(\alpha(S); S) = \sum_j \alpha(j; S) f_u(j) + e(j; \alpha(S))g + F(S)S$, where $F(S)$ is defined as the matrix of transition probabilities $F^U(\alpha(S))$. Then

$$\frac{\partial T(S)}{\partial S^0} = I - \frac{\partial \mathcal{C}(S)}{\partial S^0} = I - \left(\frac{\partial H(\alpha(S); S)}{\partial \alpha(S)} \frac{\partial \alpha(S)}{\partial S^0} + \frac{\partial H(\alpha(S); S)}{\partial S^0} \right)$$

But $\frac{\partial H(\alpha(S); S)}{\partial S^0} = F(S)$, and $\frac{\partial H(\alpha(S); S)}{\partial \alpha(S)} = 0$ by Properties (c.1) and (c.2) since $\frac{\partial H_m(\alpha(S); S)}{\partial \alpha_m} = \sum_i Q_i^{-1}(\alpha_m(S)) + \Phi v_m$ and Φv_m is always $Q_i^{-1}(\alpha_m(S))$. Therefore, $\frac{\partial T(S)}{\partial S^0} = I - F(S)$ and Newton iterations take the form $S^{K+1} = S^K + \mathcal{H}^{-1} T(S^K) = S^K + \sum_j \alpha(j; S^K) f_u(j) + e(j; \alpha(S^K))g - F(S^K)S^K = \mathcal{A}(\alpha(S^K))$. So Newton's algorithm for this problem consists of iterative application of the same policy improvement and policy valuation operators α and \mathcal{A} which define α .

Proof of (iv): The Jacobian matrices of $\mathcal{A}(\cdot)$ and $\alpha(\cdot)$ are zero at the fixed point.

By definition,

$$\mathcal{A}(\cdot) = \sum_{j=1}^n \alpha(j; \cdot) f_u(j) + e(j; \cdot)g + F(j; \cdot) \mathcal{A}(\cdot)$$

Therefore, $\mathcal{A}(\cdot) = H(\cdot; \mathcal{A}(\cdot))$. Differentiating on both sides with respect to α_m and collecting terms, one can show that

$$\frac{\partial \mathcal{A}(\cdot)}{\partial \alpha_m^0} = \mathcal{H}^{-1} \left(-F^U(\cdot) \right) \frac{\partial H(\cdot; \mathcal{A}(\cdot))}{\partial \alpha_m^0}$$

By Property (c.2) $\frac{\partial H_m}{\partial \alpha_m} = \sum_i Q_i^{-1}(\alpha_m) + \Phi v_m$. But here $\Phi v_m = Q_i^{-1}(\alpha_m)$. So at the fixed point $\Phi v_m = Q_i^{-1}(\alpha_m)$ and $\frac{\partial H_m}{\partial \alpha_m^0} = 0$. Since $\frac{\partial H_n}{\partial \alpha_m^0} = 0$ for $n \neq m$

m, we get $\frac{\partial \ln L(\mu)}{\partial \mu} = 0$. Therefore, $\frac{\partial \ln L(\mu)}{\partial \mu} = 0 = \left(\frac{\partial \ln L(\mu)}{\partial \mu} \right) \frac{\partial \mu}{\partial \mu} = 0$ at the fixed point.

Proof of Proposition 2.

[1] Since $P(\mu)$ is the unique fixed point of the mapping $\mu \rightarrow P(\mu)$, we can apply the implicit function theorem and Proposition 1 (iv) to obtain:

$$\frac{\partial P(\mu)}{\partial \mu} = \left[\frac{\partial \ln L(\mu; P(\mu))}{\partial \mu} \right]^{-1} \frac{\partial \ln L(\mu; P(\mu))}{\partial \mu} = \frac{\partial \ln L(\mu; P(\mu))}{\partial \mu} \quad (\text{Ap. 3})$$

Let μ^0 be a root in the first order conditions for the maximization of the partial likelihood function (i.e., likelihood equations):

$$\sum_{i=1}^n \frac{1}{P(d_i; x_i; \mu; \beta_f)} \frac{\partial P(d_i; x_i; \mu; \beta_f)}{\partial \mu} = 0 \quad (\text{Ap.4})$$

where $P(\mu; \beta_f) = P(\mu; \beta_f; P(\mu; \beta_f))$. Using the expression for the jacobian $\frac{\partial P(\mu)}{\partial \mu}$ in (15) we can write these likelihood equations as

$$\sum_{i=1}^n \frac{1}{P(d_i; x_i; \mu; \beta_f; P(\mu; \beta_f))} \frac{\partial P(d_i; x_i; \mu; \beta_f; P(\mu; \beta_f))}{\partial \mu} = 0 \quad (\text{Ap.5})$$

Now suppose the sequence μ^k from the NPL algorithm converges to a vector μ_{NPL} . By the Theorem of the maximum, the mapping (10) which defines the pseudo-maximum likelihood estimator is continuous in μ . Therefore, the sequence μ^k converges. Let μ_{NPL} be the limit. Then μ_{NPL} maximizes the pseudo log-likelihood function $\sum_{i=1}^n \ln P(d_i; x_i; \mu; \beta_f; \mu_{NPL})$. At μ_{NPL} the following marginal conditions of optimality hold:

$$\sum_{i=1}^n \frac{1}{P(d_i; x_i; \mu_{NPL}; \beta_f; \mu_{NPL})} \frac{\partial P(d_i; x_i; \mu_{NPL}; \beta_f; \mu_{NPL})}{\partial \mu} = 0 \quad (\text{Ap.6})$$

But convergence of the sequences on both sides of eq. (11) implies that μ_{NPL} is the fixed point of the policy iteration mapping for $\mu = \mu_{NPL}; \beta_f^0$: i.e., $\mu_{NPL} = P(\mu_{NPL}; \beta_f)$. It is clear that μ_{NPL} is a root of the likelihood equations.

[2] The same result also holds in the context of full likelihood estimation. The parameter vector is now $\mu^0 = (\mu^0; \mu_f^0)$. The term $\frac{\partial \ln L(\mu)}{\partial \mu} = \left(\frac{\partial \ln L(\mu)}{\partial \mu}; \frac{\partial \ln L(\mu)}{\partial \mu_f} \right)$ is

added to the first order conditions. Since exactly the same term is added to both likelihood and pseudo-likelihood equations, intuitively it should not have any effect on equivalence. More explicitly, the likelihood equations are

$$\sum_{i=1}^n \frac{1}{a(d_i; x_i; \beta; P[\beta])} \frac{\partial^a (d_i; x_i; \hat{\mu}; P[\beta])}{\partial \mu} + \frac{\partial l_2(\beta)}{\partial \mu} = 0 \quad (\text{Ap.7})$$

which are again the first order conditions satisfied by NPL's limit pair $(\mu_{\text{NPL}}; \beta_{\text{NPL}})$ in a full likelihood context.

Proof of Proposition 3.

Let $\hat{A}_n(\cdot) \equiv a(\theta_n(\cdot); \beta_{fn}; \cdot)$ be the NPL operator, where $\theta_n(\cdot) \equiv \arg \max_{\theta} \ell_n(\theta; \beta_{fn}; \cdot)$ and $\ell_n(\theta; \beta_{fn}; \cdot)$ is the pseudo-likelihood given \cdot and a sample of size n . Let θ_n be a root of the likelihood equations, and $\hat{P}_n \equiv P(\theta_n; \beta_{fn})$. It is clear from the proof of Proposition 2 that \hat{P}_n is a fixed point of $\hat{A}_n(\cdot)$ and $\theta_n = \theta_n(\hat{P}_n)$. In order to establish local convergence of fixed point iterations, we first obtain the Jacobian of $\hat{A}_n(\cdot)$ at the fixed point \hat{P}_n :

$$\frac{\partial \hat{A}_n(\beta_n)}{\partial \beta_n} = \frac{\partial^a (\theta_n; \beta_{fn}; \beta_n)}{\partial \mu^0} \frac{\partial \theta_n(\beta_n)}{\partial \beta_n} + \frac{\partial^a (\theta_n; \beta_{fn}; \beta_n)}{\partial \beta_n} = \frac{\partial^a (\theta_n; \beta_{fn}; \beta_n)}{\partial \mu^0} \frac{\partial \theta_n(\beta_n)}{\partial \beta_n}$$

since the second term is zero at all samples by Proposition 1. We now differentiate the pseudo-likelihood equations $\ell_n(\theta_n; \beta_{fn}; \beta_n) = 0$ in order to obtain $\frac{\partial \theta_n}{\partial \beta_n}$:

$$\frac{\partial \theta_n(\beta_n)}{\partial \beta_n} = - \frac{\partial^2 \ell_n(\theta_n; \beta_{fn}; \beta_n)}{\partial \mu \partial \beta_n} \#_{i=1}^{-1} \frac{\partial^2 \ell_n(\theta_n; \beta_{fn}; \beta_n)}{\partial \mu \partial \beta_n}$$

In part (B.1) of the proof of Proposition 4 we show that $(1/n) \partial^2 \ell_n = o_p(\mu^0)$ and $(1/n) \partial^2 \ell_n = o_p(\mu^0)$ if ℓ_n is evaluated at consistent estimators. The root of the likelihood equations θ_n is consistent under the regularity conditions we have imposed, and β_{fn} is also consistent by assumption, so $\beta_n \xrightarrow{p} P^*$. Furthermore, $\frac{\partial^a (\theta_n; \hat{\mu}_{fn}; \beta_n) = o_p(\mu^0) \xrightarrow{p} \partial P(\mu^*) = o_p(\mu^0)$ by Proposition 1 (iv) and the Slutsky theorem. Therefore, $\frac{\partial \hat{A}_n(\beta_n) = o_p(\mu^0) \xrightarrow{p} 0$. Let λ_n denote the spectral radius of $\frac{\partial \hat{A}_n(\beta_n) = o_p(\mu^0)$. Then $\lambda_n \xrightarrow{p} 0$ by the Slutsky theorem since the spectral radius is continuous. So for

any \pm ; $0 < \pm < 1$, we have $\Pr(\|\hat{A}_n - A\| < \pm) \rightarrow 1$. But fixed point iterations converge locally if $\|\hat{A}_n\| < 1$ by Theorem 5.4.2 in Judd (1998) since $\hat{A}_n(\cdot)$ is Lipschitz.

Proof of Proposition 4

Let n denote sample size. Consider the functions:

$$Q_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ln^a(d_i; x_i; \theta; \mu_f; \cdot) \quad ; \quad Q_n^K(\theta) = Q_n(\theta; \hat{\mu}_{fn}; \hat{1}_n^{K_i-1});$$

(A) Consistency of θ^K

The proof of consistency proceeds in stages:

(A.1) If $\hat{1}_n^{K_i-1}$ and $\hat{\mu}_{fn}$ are consistent, then $Q_n^K(\theta)$ converges a.s. and uniformly in θ to a function $Q_1(\theta)$.

(A.2) If $\hat{1}_n^{K_i-1}$ and $\hat{\mu}_{fn}$ are consistent, then $\theta_n^K = \arg \max_{\theta \in \Theta} Q_n^K(\theta)$ converges a.s. to θ^K .

(A.3) For $K \geq 1$, if $\hat{1}_n^{K_i-1} \xrightarrow{a.s.} 1$ and $\hat{\mu}_{fn} \xrightarrow{a.s.} \mu_f^a$, then $\theta_n^K \xrightarrow{a.s.} \theta^K$.

By condition (viii), $\hat{\mu}_{fn} \xrightarrow{a.s.} \mu_f^a$ and $\hat{1}_n^0 \xrightarrow{a.s.} 1$; a simple induction argument based on (A.2) and (A.3) completes the proof.

(A.1) Define $Q_1(\theta) = E(\ln^a[d_i; x_i; \theta])$. By Lemma 24.1 in Gourieroux and Monfort (vol. II, page 392), we have that if: (a) $Q_n(\theta)$ converges a.s. and uniformly in θ to $Q_1(\theta)$; (b) $Q_1(\theta)$ is uniformly continuous in θ ; and (c) $f_{\hat{\mu}_{fn}}g$ and $f_{\hat{1}_n^{K_i-1}}g$ converge a.s. to μ_f^a and 1 , respectively; then $Q_n(\theta; \hat{\mu}_{fn}; \hat{1}_n^{K_i-1})$ converges a.s. and uniformly in θ to $Q_1(\theta; \mu_f^a; 1) = Q_1(\theta)$. We prove that conditions (a), (b) and (c) hold.

Let $H_x(\cdot; \cdot)$ be the true probability distribution of $(d; x)$, and let $H_n(\cdot; \cdot)$ be the empirical distribution of $(d; x)$ in a sample with size n . By definition, for any θ :

$$\begin{aligned} |Q_n(\theta) - Q_1(\theta)| &= \left| \frac{1}{n} \sum_{j=1}^n \ln^a(j; x; \theta) [H_n(j; x) - H_x(j; x)] \right| \\ &\leq \frac{1}{n} \sum_{j=1}^n |\ln^a(j; x; \theta)| |H_n(j; x) - H_x(j; x)| \end{aligned}$$

By condition (iv) $H_n(j; x) \xrightarrow{a.s.} H_\alpha(j; x)$. Furthermore, by conditions (i)-(iii) $\ln^\alpha(j; x; \theta)$ is bounded. Therefore,

$$\begin{aligned} & \Pr \limsup_{n \rightarrow \infty} \sup_{\theta \in \Theta} |Q_n(\theta) - Q_1(\theta)| = 0 \\ & \Pr \limsup_{n \rightarrow \infty} \sup_{\theta \in \Theta} |\ln^\alpha(j; x; \theta) - \ln^\alpha(j; x; \theta)| |H_n(j; x) - H_\alpha(j; x)| = 0 \\ & \Pr \lim_{n \rightarrow \infty} |H_n(j; x) - H_\alpha(j; x)| = 0 \end{aligned}$$

i.e., $Q_n(\theta)$ converges a.s. and uniformly in θ to $Q_1(\theta)$.

By conditions (i) and (ii), $Q_1(\theta)$ is continuous on a compact set, so it is uniformly continuous, i.e., (b) holds. Finally, condition (viii) says that $\hat{\mu}_{fn} \xrightarrow{a.s.} \mu_f^\alpha$, and $\hat{\mu}_n \xrightarrow{a.s.} \mu^\alpha$ by assumption.

(A.2) By Property 24.2 in Gourieroux and Monfort (vol. II, page 387), if: (a) $Q_n^K(\theta)$ converges a.s. and uniformly in θ to $Q_1^K(\theta)$; and (b) $Q_1^K(\theta)$ has a unique maximum in \mathbb{R}^m at θ^α ; then $\hat{\theta}_n^K \xrightarrow{a.s.} \theta^\alpha$. We have proved (a) in (A.1). Condition (vi) implies (b) by the information inequality.

(A.3) By definition $\hat{\theta}_n^K = \hat{\theta}_n^K(\hat{\mu}_{fn}; \hat{\mu}_n^{K-1})$. By (A.2) $\hat{\theta}_n^K \xrightarrow{a.s.} \theta^\alpha$. Since $\hat{\theta}_n^K(\cdot)$ is continuous in θ , by the Slutsky theorem $\hat{\theta}_n^K \xrightarrow{a.s.} \theta^\alpha(\mu_f^\alpha; \mu^\alpha) = \theta^\alpha$.

(B) Asymptotic distribution of $\hat{\theta}_n^K$

The proof is based on the following two results:

(B.1) Let $\hat{\mu}_n$ be a consistent estimator of μ^α , and define $\hat{\theta}_n^K = \hat{\theta}_n^K(\hat{\mu}_{fn}; \hat{\mu}_n^{K-1})$. If $\sqrt{n}(\hat{\mu}_{fn} - \mu_f^\alpha) \xrightarrow{d} N(0; V_f^\alpha)$; $\sqrt{n}(\hat{\mu}_n - \mu^\alpha) \xrightarrow{d} N(0; V^\alpha)$; $\sqrt{n}(\hat{\theta}_n^K - \theta^\alpha) \xrightarrow{d} N(0; V^\alpha)$, where V^α does not depend on n .

(B.2) If $\sqrt{n}(\hat{\mu}_{fn} - \mu_f^\alpha) \xrightarrow{d} N(0; V_f^\alpha)$; $\sqrt{n}(\hat{\mu}_n^{K-1} - \mu^\alpha) \xrightarrow{d} N(0; V^\alpha)$; $\sqrt{n}(\hat{\theta}_n^K - \theta^\alpha) \xrightarrow{d} N(0; V^\alpha)$ for $K \geq 1$. It follows that the upper left $r \times r$ submatrices of V and V^α are identical, where r is the dimension of the parameter vector $\theta^\alpha; \mu_f^\alpha$.

The proof is completed by induction using (B.1) and (B.2) and condition (viii).

(B.1) Given conditions (ii) and (vii) and the definition of $\hat{\theta}_n^K$, the first order conditions of optimality imply that with probability approaching one $\hat{\theta}_n^K(\hat{\mu}_{fn}) = \theta^\alpha = 0$,

where $\theta_n \sim (\theta_n; \mu_{fn}; \hat{\theta}_n)$. By condition (ii), $Q_n(\cdot)$ is twice continuously differentiable and we can apply the stochastic mean value theorem to $Q_n(\cdot)$ between θ_n and θ^* . There are p vectors $\eta_n^1; \eta_n^2; \dots; \eta_n^p$ which are convex combinations of θ_n and θ^* such that:

$$\frac{\partial Q_n(\theta_n)}{\partial \theta} \Big|_i - \frac{\partial Q_n(\theta^*)}{\partial \theta} = \begin{matrix} \mathbf{O} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{A} \end{matrix} \begin{matrix} \partial^2 Q_n(\eta_n^1) \\ \vdots \\ \partial^2 Q_n(\eta_n^p) \end{matrix} \begin{matrix} \mathbf{1} \\ \mathbf{A}(\theta_n; \theta^*) \end{matrix}$$

Given that any η_n^j is a convex linear combination of θ_n and θ^* , and given that $\theta_n \xrightarrow{a.s.} \theta^*$, then $\eta_n^j \xrightarrow{a.s.} \theta^*$. Furthermore, $\partial^2 Q_n(\theta^*)$ converges in probability and uniformly in θ to $\partial^2 Q_1(\theta^*)$ and therefore $\partial^2 Q_n(\eta_n^j) \xrightarrow{p} \partial^2 Q_1(\theta^*)$ (see Amemiya, Thm 4.2.2 and Thm.4.1.5). We can now rewrite the previous mean value theorem as follows:

$$\begin{aligned} \frac{\partial Q_n(\theta_n)}{\partial \theta} \Big|_i &= \frac{\tilde{\mathbf{A}} \partial^2 Q_1(\theta^*)}{\partial \theta \partial \theta} + o_p(1) \Big|_{(\theta_n; \theta^*)} + \frac{\tilde{\mathbf{A}} \partial^2 Q_1(\theta^*)}{\partial \theta \partial \mu_f} + o_p(1) \Big|_{(\hat{\mu}_{fn}; \mu_f^*)} \\ &+ \frac{\tilde{\mathbf{A}} \partial^2 Q_1(\theta^*)}{\partial \theta \partial \theta} + o_p(1) \Big|_{(\hat{\theta}_n; P^*)} \end{aligned}$$

Or:

$$\begin{aligned} P_{\tilde{n}(\theta_n; \theta^*)} &= \begin{matrix} \tilde{\mathbf{A}} \\ \mathbf{I} \end{matrix} \frac{\partial^2 Q_1(\theta^*)}{\partial \theta \partial \theta} + o_p(1) \\ &\begin{matrix} \tilde{\mathbf{A}} \\ \mathbf{I} \end{matrix} \frac{\partial^2 Q_1(\theta^*)}{\partial \theta \partial \mu_f} + o_p(1); \begin{matrix} \mathbf{O} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{A} \end{matrix} \begin{matrix} P_{\tilde{n}(\theta_n; \theta^*)} \\ P_{\tilde{n}(\hat{\mu}_{fn}; \mu_f^*)} \\ P_{\tilde{n}(\hat{\theta}_n; P^*)} \end{matrix} \end{aligned}$$

Notice that condition (vi) implies that $\partial^2 Q_1(\theta^*)$ is a non singular (negative definite) matrix. By condition (viii) and the Mann-Wald Theorem, it is straightforward that $P_{\tilde{n}(\theta_n; \theta^*)} \xrightarrow{d} N(0; V^*)$, where:

$$V^* = \begin{matrix} \tilde{\mathbf{A}} \\ \mathbf{I} \end{matrix} \frac{\partial^2 Q_1(\theta^*)}{\partial \theta \partial \theta} \begin{matrix} \tilde{\mathbf{A}} \\ \mathbf{I} \end{matrix} + \begin{matrix} \tilde{\mathbf{A}} \\ \mathbf{I} \end{matrix} \frac{\partial^2 Q_1(\theta^*)}{\partial \theta \partial \mu_f} \begin{matrix} \tilde{\mathbf{A}} \\ \mathbf{I} \end{matrix} + \begin{matrix} \mathbf{O} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{A} \end{matrix} \begin{matrix} P_{\tilde{n}(\theta_n; \theta^*)} \\ P_{\tilde{n}(\hat{\mu}_{fn}; \mu_f^*)} \\ P_{\tilde{n}(\hat{\theta}_n; P^*)} \end{matrix}$$

Using Proposition 1 it is possible to simplify this expression. First, notice that,

$$\frac{\partial^2 Q_1(\theta^*)}{\partial \theta \partial \theta} = E \frac{\partial^2 \ln^a(d_i; x_i; \theta^*)}{\partial \theta \partial \theta} = E \frac{\partial \ln^a(d_i; x_i; \theta^*)}{\partial \theta} \frac{\partial \ln^a(d_i; x_i; \theta^*)}{\partial \theta}$$

APPENDIX 2: Description of the algorithms

NFXP algorithm

[1] Inner (fixed point) algorithm:

Smooth policy iteration algorithm. with convergence criterion $\max_{f_j, mg} \sum_{i,j,m} |f_{i,j,m}^k - f_{i,j,m}^{k-1}| < 10^{-6}$. At any outer iteration k , the initial guess for the vector of probabilities is $P^{(k-1)}$. These "smart guesses" reduce very significantly the number of policy iterations as we proceed with the outer iterations.

[2] Outer (hill climbing) algorithm:

BHHH method with convergence criterion $\max_{f_j, g} \sum_j |f_j^{(k)} - f_j^{(k-1)}| < 10^{-6}$. We use an analytical expression for the gradient of the likelihood function (the gradient is a by-product of a policy iteration) and therefore we have to solve one time the fixed point problem for each outer iteration. τ is fixed at 0.9999.

NPL algorithm

[1] Inner (pseudo-ML) algorithm:

It is a simple Multinomial logit estimation. The convergence criterion is $\max_{f_j, g} \sum_j |f_j^{(k)} - f_j^{(k-1)}| < 10^{-6}$. τ is fixed at 0.9999.

[2] Outer algorithm:

It is a "Pseudo" Policy iteration procedure. It is "pseudo" because at each iteration we use a new value of θ to obtain the new vector of probabilities. The convergence criterion is $\max_{f_j, mg} \sum_{i,j,m} |f_{i,j,m}^k - f_{i,j,m}^{k-1}| < 10^{-6}$.

REFERENCES:

Bierens, H. (1983): "Uniform consistency of kernel estimators of a regression function under generalized conditions", *Journal of the American Statistical Association*, 78, 669-707.

Delgado, M. and J. Mora (1995a): "Non parametric and semiparametric estimation with discrete regressors", *Econometrica* 63 (6), pp. 1477-1484.

Delgado, M. and J. Mora (1995b): "On asymptotic inferences in non parametric and semiparametric models with discrete and mixed regressors", *Investigaciones Economicas XIX* (3), pp. 435-468.

Eckstein, Z., and K. Wolpin (1989): "The specification and estimation of dynamic stochastic discrete choice models", *Journal of Human Resources* 24, pp. 562-598.

Gourieroux, C., and A. Monfort (1995): "Statistics and Econometric Models: Volume I", Cambridge University Press.

Hotz, J., and R.A. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models", *Review of Economic Studies* 60, pp. 497-529.

Hotz, J., R.A. Miller, S. Sanders, and J. Smith (1994): "A simulation estimator for dynamic models of discrete choice", *Review of Economic Studies* 61, pp. 265-89.

Judd, K. (1998): "Numerical Methods in Economics", MIT Press.

Keane, M. P., and K. I. Wolpin (1994): "The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence", *The Review of Economics and Statistics* 76, pp. 648-672.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men", *Journal of Political Economy*, 105, pp. 473-522.

Manski, C. (1991): "Nonparametric estimation of expectations in the analysis of discrete choice under uncertainty", in *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, W. Barnett, J. Powell, and G. Tauchen (eds.). Cambridge University Press.

Manski, C. (1993): "Dynamic choice in social settings: Learning from the experience of others", *Journal of Econometrics* 58, pp. 121-36.

McFadden, D. (1973): "Conditional logit analysis of qualitative choice behavior",

in P. Zarembka (ed.) *Frontiers of Econometrics*, New York, Academic Press, pp. 105-42.

Miller, R. A. (1997): "Estimating Models of Dynamic Optimization with Microeconomic Data", in H. Pesaran and P. Smidh (eds.) *Handbook of Applied Econometrics: Microeconomics (Vol 2)*. Blackwell.

Pakes, A, and P. McGuire (1997): "Stochastic Algorithms for Dynamic Models: Markov Perfect Equilibrium and the 'Curse of Dimensionality'", manuscript, Department of Economics, Yale University.

Pakes, A (1994): "Dynamic structural models, problems and prospects", in J. J. LaFont and C. Sims (eds.), *Advances in Econometrics, Proceedings of the Sixth World Congress of the Econometric Society*. Cambridge University Press.

Puterman, M. (1994): "Markov Decision Processes", Wiley, New York.

Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher", *Econometrica* 55, pp. 999-1033.

Rust, J. (1988): "Maximum Likelihood Estimation of Discrete Control Processes", *SIAM Journal of Control and Optimization*, pp. 1006-1024

Rust, J. (1994): "Structural estimation of Markov decision processes", in R. E. Engle and McFadden (eds.) *Handbook of Econometrics Volume 4*, North-Holland.

Rust, J. (1996): "Numerical Dynamic Programming in Economics", in H Amman, D. Kendrick, and J. Rust (eds.). *Handbook of Computational Economics*, Elsevier-North Holland, Chapter 14.

Rust, J. (1997a): "Using Randomization to Break the Curse of Dimensionality", *Econometrica*, 65, pp. 487-516.

Rust, J. (1997b): "A Comparison of Policy Iteration Methods for Solving Continuous-State, Infinite-Horizon Markovian Decision Problems Using Random, Quasi-Random, and Deterministic Discretizations", manuscript, Department of Economics, Yale University.

Table 1
Monte Carlo Experiment 1
Sample size = 10,000; Initial probabilities = Kernel estimates

Without trimming				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	29.8	0.4	0.2
	Median absolute error (% over Median AE of MLE)	19.8	1.0	0.6
	Std. dev. estimator (% over Std. dev. of MLE)	29.4	0.4	0.1
μ_1	Mean absolute error (% over Mean AE of MLE)	44.4	0.9	0.2
	Median absolute error (% over Median AE of MLE)	21.6	0.7	0.5
	Std. dev. of estimator (% over Std. dev. of MLE)	50.1	1.1	0.2

With trimming: Percentiles 5 to 95				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	26.0	0.5	0.2
	Median absolute error (% over Median AE of MLE)	19.8	1.0	0.6
	Std. dev. estimator (% over Std. dev. of MLE)	22.0	0.5	0.2
μ_1	Mean absolute error (% over Mean AE of MLE)	36.0	1.0	0.2
	Median absolute error (% over Median AE of MLE)	21.6	0.7	0.5
	Std. dev. estimator (% over Std. dev. of MLE)	36.0	1.1	0.2

Average number of Policy Iterations to obtain MLE = 5.3

True parameters: $\mu_0 = 10.47$; $\mu_1 = 0.58$; $\bar{\pi} = 0.9999$
State space = 201 ; Number of replications = 1000
All entries are $100 * (\text{K-PI statistic} - \text{MLE statistic}) / \text{MLE statistic}$

Table 2
Monte Carlo Experiment 2
Sample size = 5,000; Initial probabilities = Kernel estimates

Without trimming				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	20.6	0.7	0.2
	Median absolute error (% over Median AE of MLE)	16.4	0.5	-0.4
	Std. dev. estimator (% over Std. dev. of MLE)	18.8	0.6	0.2
μ_1	Mean absolute error (% over Mean AE of MLE)	35.5	0.7	0.3
	Median absolute error (% over Median AE of MLE)	28.8	0.6	0.5
	Std. dev. of estimator (% over Std. dev. of MLE)	36.4	0.8	0.2

With trimming: Percentiles 5 to 95				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	18.6	0.6	0.2
	Median absolute error (% over Median AE of MLE)	16.4	0.5	-0.4
	Std. dev. estimator (% over Std. dev. of MLE)	14.4	0.7	0.1
μ_1	Mean absolute error (% over Mean AE of MLE)	29.7	0.6	0.3
	Median absolute error (% over Median AE of MLE)	28.8	0.6	0.5
	Std. dev. estimator (% over Std. dev. of MLE)	27.7	0.8	0.2

Average number of Policy Iterations to obtain MLE = 5.5

True parameters: $\mu_0 = 10.47$; $\mu_1 = 0.58$; $\bar{\pi} = 0.9999$
State space = 201 ; Number of replications = 1000
All entries are 100*(K-PI statistic - MLE statistic)/MLE statistic

Table 3
Monte Carlo Experiment 3
Sample size = 1,000; Initial probabilities = Kernel estimates

Without trimming				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	4.7	1.6	0.3
	Median absolute error (% over Median AE of MLE)	14.2	0.2	-0.3
	Std. dev. estimator (% over Std. dev. of MLE)	6.8	1.2	0.3
μ_1	Mean absolute error (% over Mean AE of MLE)	18.7	1.5	0.2
	Median absolute error (% over Median AE of MLE)	25.1	0.7	0.6
	Std. dev. of estimator (% over Std. dev. of MLE)	11.0	1.3	0.2

With trimming: Percentiles 5 to 95				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	8.1	1.7	0.3
	Median absolute error (% over Median AE of MLE)	14.2	0.2	-0.3
	Std. dev. estimator (% over Std. dev. of MLE)	6.9	1.5	0.3
μ_1	Mean absolute error (% over Mean AE of MLE)	22.7	1.6	0.2
	Median absolute error (% over Median AE of MLE)	25.1	0.7	0.6
	Std. dev. estimator (% over Std. dev. of MLE)	11.4	1.2	0.2

Average number of Policy Iterations to obtain MLE = 6.2

True parameters: $\mu_0 = 10.47$; $\mu_1 = 0.58$; $\bar{\pi} = 0.9999$
State space = 201 ; Number of replications = 1000
All entries are $100 * (\text{K-PI statistic} - \text{MLE statistic}) / \text{MLE statistic}$

Table 4				
Monte Carlo Experiment 4				
Sample size = 10,000 ; Initial probabilities = Kernel + Noise				
Without trimming				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	53.1	2.1	0.0
	Median absolute error (% over Median AE of MLE)	41.8	2.0	0.7
	Std. dev. estimator (% over Std. dev.of MLE)	34.8	1.8	-0.2
μ_1	Mean absolute error (% over Mean AE of MLE)	127.8	2.8	0.0
	Median absolute error (% over Median AE of MLE)	117.3	1.2	0.1
	Std. dev. of estimator (% over Std. dev.of MLE)	59.3	2.6	-0.1
With trimming: Percentiles 5 to 95				
Parameter	Statistics	Estimators		
		1-PI	2-PI	3-PI
μ_0	Mean absolute error (% over Mean AE of MLE)	47.4	1.8	0.0
	Median absolute error (% over Median AE of MLE)	41.8	2.0	0.7
	Std. dev. estimator (% over Std. dev.of MLE)	25.2	1.2	-0.1
μ_1	Mean absolute error (% over Mean AE of MLE)	120.6	2.4	0.0
	Median absolute error (% over Median AE of MLE)	117.3	1.2	0.1
	Std. dev. estimator (% over Std. dev.of MLE)	41.5	1.7	-0.1

Average number of Policy Iterations to obtain MLE = 5.7

True parameters: $\mu_0 = 10.47$; $\mu_1 = 0.58$; $\bar{\pi} = 0.9999$

State space = 201 ; Number of replications = 1000

All entries are $100 * (\text{K-PI statistic} - \text{MLE statistic}) / \text{MLE statistic}$

Table 5
Estimated standard errors and Monte Carlo distribution

Experiment 1: N=10,000; Kernel					
Parameter	Statistics	Estimators			
		1-PI	2-PI	3-PI	MLE
μ_0	Ratio ^(a)	0.719	1.034	1.038	1.035
μ_1	Ratio ^(a)	0.598	1.048	1.063	1.056

Experiment 2: N=5,000; Kernel					
Parameter	Statistics	Estimators			
		1-PI	2-PI	3-PI	MLE
μ_0	Ratio ^(a)	0.791	1.073	1.080	1.075
μ_1	Ratio ^(a)	0.643	1.073	1.087	1.079

Experiment 3: N=1,000; Kernel					
Parameters	Statistics	Estimators			
		1-PI	2-PI	3-PI	MLE
μ_0	Ratio ^(a)	0.801	1.008	1.027	1.022
μ_1	Ratio ^(a)	0.666	1.043	1.076	1.065

Experiment 4: N=10,000; Kernel + Noise					
Parameter	Statistics	Estimators			
		1-PI	2-PI	3-PI	MLE
μ_0	Ratio ^(a)	0.686	1.041	1.066	1.059
μ_1	Ratio ^(a)	0.549	1.044	1.077	1.067

Note: Ratio between the average estimated asymptotic standard error (over the 1000 replications) and the empirical standard deviation of the estimator (over the 1000 replications).





