

Federal Reserve Bank of Minneapolis
Research Department

**Bayesian Inference for Dynamic
Discrete Choice Models Without
the Need for Dynamic Programming**

John F. Geweke and Michael P. Keane*

Working Paper 564

January 1996

*Both, Federal Reserve Bank of Minneapolis and University of Minnesota. Keane's work on this project was supported by a grant from the National Science Foundation - NSF/SBR 9511186. Dan House provided excellent research assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

I. Introduction

This paper presents a method for Bayesian inference with regard to the structural parameters of dynamic discrete choice models. The method does not require the econometrician to solve the agents' optimization problem, or make strong assumptions about how agents form expectations. It is implemented using a Gibbs sampling-Data augmentation algorithm (Gelfand and Smith (1990), Tanner and Wong (1987)). We first describe the generic discrete choice model framework, describe our method for Bayesian inference, and then present some results obtained by applying the method to trial problems.

We consider models in which agents choose from among a set of J mutually exclusive alternatives, and make choices repeatedly over T periods. Such models can usually be written in the form:

$$(1) \quad V_{jt}(S_i(t)) = P_{jt}(S_i(t)) + \delta E_t[V(S_i(t+1)) | S_i(t), d_{jt} = 1]$$

and

$$(2) \quad S_i(t+1) = M[S_i(t), d_{jt}].$$

where $S_i(t)$ is the state of agent i at time t , $V_{jt}(S_i(t))$ is the value to agent i of choosing alternative j at time t , $P_{jt}(S_i(t))$ is the time t payoff to agent i from choosing alternative j , δ is the discount factor, E_t is the expectation operator, $V(S_i(t+1))$ is the value of being in state $S_i(t+1)$, d_{jt} is an indicator equal to 1 if alternative j is chosen at time t and zero otherwise, d_{jt} is the $J \times 1$ vector of the J choice indicators, and M is a random function that maps $S_i(t)$ and d_{jt} into $S_i(t+1)$.

Equation (1) says that the value of choosing alternative j at time t is the immediate payoff plus a "future component" equal to the discount factor times the expected value of the state the agent attains at time $t+1$ given the current state $S_i(t)$ and current choice. Since $S_i(t+1)$ is determined by the random function $M[S_i(t), d_{jt}]$, the expectation is taken over realizations from M .

We consider situations in which the econometrician is willing to assume parametric functional forms for the payoff functions $P_{jt}(S_i(t))$ and the law of motion for the state variables (equation (2)), and in which the econometrician observes at least some information about payoffs (e.g., the value of a payoff $P_{jt}(S_i(t))$ may be observed if and only if $d_{jt} = 1$).

The econometrician is interested in learning about the structural parameters that enter the payoff functions. However, the econometrician is either unwilling to make strong assumptions about the process by which expectations are formed, or lacks the necessary computational power to calculate the $E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$ functions given such assumptions. As an example of the latter case, suppose that a rational expectations assumption is invoked. Then, solving for the future components of the value functions requires that the agents' optimization problem be solved via dynamic programming. This will be an extremely computationally burdensome process for all but the most simple model structures (See Keane and Wolpin (1994) or Rust (1995) for a discussion).

If it were feasible to solve for the $E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$ functions rather quickly via dynamic programming, then classical inference with regard to the structural parameters would be feasible. But, even then, Bayesian inference would likely remain infeasible. This is because the $E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$ functions are themselves complex nonlinear functions of the deep structural parameters. Thus, the mapping from the observed data (choices and payoffs) to the posterior distribution of the structural parameters will typically be intractably complex. (Recently, Lancaster (1995) has shown that the mapping is quite tractable in the simple infinite horizon job search model, but this special case is unusual.)

The idea behind our approach to Bayesian inference for dynamic discrete models is that much can be learned about the future components of the value functions just by looking at observed choices and payoffs - without the need to assume an expectations formation mechanism or solve agents' optimization problem. Suppose that agents in a particular state usually choose alternative 1, despite the fact that their observed payoffs are relatively low as compared to those of other agents in the same state who choose alternative 2. This suggests that the future component of the value function associated with choice of alternative 1 is greater than that associated with 2 (e.g., the fact that young people often choose to go to school, despite the fact that immediate payoffs in terms of income would be greater if they chose to work, suggests that future component of the value function associated with the school choice is greater than that associated with the work choice for many young people). Of course, if payoffs are only partially observable, such inferences about the future components of the value functions will hinge on functional form assumptions for the payoff functions.

Following the above intuition, our approach is to specify a flexible functional form for the future component of the value function $\delta E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$, expressing it as a polynomial in the state variables. Then, given assumed functional forms for the payoff functions, and given some information about actual payoffs, it is possible to learn both about the

parameters of the payoff functions and the structure of expectations. Using a Gibbs sampling-data augmentation algorithm, it is straightforward to simulate the joint posterior distribution of the structural parameters of the payoff functions and the parameters characterizing expectations. We illustrate this approach in Section II.

It is worth noting that our approach is somewhat related to that of Hotz and Miller (1993), who also develop a method for estimating dynamic discrete choice models without the need for dynamic programming. If agents solve a dynamic programming problem to arrive at their decision rules, then $E_t[V(S_i(t+1)) | S_i(t), d_{ijt}=1]$ in (1) takes the form:

$$(3) \quad E_t[V(S_i(t+1)) | S_i(t), d_{ijt}=1] = E_t[\max_k \{V_{ikt}(S_i(t+1)) | S_i(t), d_{ijt}=1\}]$$

where the expectation is taken over realizations for $S_i(t+1)$ determined by the random function M . Given the definition in (3), equation (1) becomes the Bellman equation (Bellman (1957)). In this case, Hotz and Miller point out that $E_t[V(S_i(t+1)) | S_i(t), d_{ijt}=1]$ can be written as a function of agents' choice probabilities in periods $t+1$ through T , conditional on every state that may be occupied from $t+1$ through T , and the current payoff functions for each alternative in those states. Thus, the future components of the value functions can be calculated, given data on choice probabilities conditional on all possible states, and assumed functional forms for the payoff functions. Given such calculations, it is not necessary to solve a dynamic programming problem to construct the future components of the value functions in (1).

In practice, of course, the econometrician will not typically have access to data on the conditional choice probabilities at every point in the state space. Suppose one has a panel data set on choices over time for a set of individuals. One can only construct the choice frequencies at the subset of state points actually observed in the data. To these one can apply nonparametric smoothing techniques to obtain estimated conditional choice probabilities at all state points. The predicted conditional choice probabilities are then used to form the $E_t[V(S_i(t+1)) | S_i(t), d_{ijt}=1]$ functions.

The Hotz-Miller procedure suffers from a number of limitations. First, the data requirements are heavy. For large state space problems, the ratio of observed choices to state points will tend to be small unless one has access to an enormous data set. Thus, as Rust (1995) has observed, the results may be very sensitive to choice of nonparametric smoothing technique. Second, a strong stationarity assumption is required. Panel data sets typically only contain data collected over a fraction of each individual's life. Therefore, the choice frequencies for periods far in the future that are used in estimating the future component of an agent's value function

will typically be the choice frequencies that were observed for much older cohorts of agents. The implicit assumption that a younger cohort's choice probabilities will be identical to those of older cohorts when they reach the same state points rules out regime changes over time. Third, unobserved heterogeneity is ruled out, since conditional choice frequencies for all agents are used to calculate the choice probabilities of all agents in the same observed state.

A related approach to the problem is that of Manski (1993), who assumes that agents form the $E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$ functions by looking at the life histories of payoffs for older agents who at some past date were in the state $S_i(t)$ and made the choice $d_{ijt} = 1$. This again entails a strong stationarity assumption.

Our approach to inference in dynamic discrete choice models has important advantages over these other approaches. It does not require the assumption of stationarity in the economic environment across cohorts, since the method could be applied to a single cohort observed over a fraction of the planning period $t=1, T$. Our approach can admit unobserved heterogeneity, provided the econometrician specifies the manner in which the unobservable affects the payoff functions. And our approach does not make heavy data requirements, other than that the payoffs be at least partially observable. We do need to place some structure on the future component of the value function $\delta E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$, although this can be a flexible functional form such as a high order polynomial in the state variables. And we do need to assume functional forms for the current payoff functions, but this is true of the other approaches as well.

II. An Illustrative Application

In this section we present an illustrative application of our method of Bayesian inference for dynamic discrete choice models. Consider a homogeneous population of agents who choose between two occupations in each of T periods. Wages in each occupation are determined by the functions:

$$(4) \quad \ln W_{i1t} = \beta_{01} + \beta_{11}X_{i1t} + \beta_{21}X_{i1t}^2 + \beta_{31}X_{i2t} + \beta_{41}X_{i2t}^2 + \epsilon_{i1t}$$

$$(5) \quad \ln W_{i2t} = \beta_{02} + \beta_{12}X_{i2t} + \beta_{22}X_{i2t}^2 + \beta_{32}X_{i1t} + \beta_{42}X_{i1t}^2 + \epsilon_{i2t}$$

where X_{jt} is experience in occupation j at time t , and:

$$(6) \quad \epsilon_{it} \equiv (\epsilon_{i1t}, \epsilon_{i2t})' \sim N(0_2, \Sigma).$$

The ϵ_{it} are assumed to be serially independent. We also assume that each alternative has associated with it a nonpecuniary benefit ν_{ijt} , where $\nu_{it} \equiv (\nu_{i1t}, \nu_{i2t})' \sim N(0_2, \Sigma)$. In this model, the state of agent i at time t is:

$$(7) \quad S_i(t) = (X_{i1t}, t, \epsilon_{it}, \nu_{it}).$$

Note that it is redundant to include the occupation 2 experience level in the state space, since $X_{i1t} + X_{i2t} = t$.

Notice that while the state variables X_{i1t} and t evolve deterministically (i.e., $X_{i1,t+1} = X_{i1t} + 1$), the state variables ϵ_{it} and ν_{it} are serially independent. Thus, the future component of the value function, $\delta E_t[V(S_i(t+1)) | S_i(t), d_{ijt} = 1]$, is a function of only $X_{i1,t+1}$ and t . In writing the value functions, we will suppress the argument $S_i(t)$ and instead subscript by i so as to keep the notation more compact.

The value functions associated with each alternative are:

$$(8) \quad V_{i1t} = W_{i1t} + \nu_{i1t} + F(X_{i1t} + 1, t)$$

$$(9) \quad V_{i2t} = W_{i2t} + \nu_{i2t} + F(X_{i1t}, t)$$

where $F(X_1, t)$ is the expected present value of arriving at time t with X_1 as the occupation 1 experience level.

Since choices depend only on the differences, rather than the levels, of the value functions, we define:

$$\begin{aligned}
 (10) \quad Z_t &\equiv V_{i1t} - V_{i2t} \\
 &= W_{i1t} - W_{i2t} + \nu_{i1t} - \nu_{i2t} + F(X_{i1t}+1, t) - F(X_{i1t}, t) \\
 &= W_{i1t} - W_{i2t} + \eta_{i1t} + f(X_{i1t}+1, t)
 \end{aligned}$$

where $\eta_{i1t} = \nu_{i1t} - \nu_{i2t} \sim N(0, \sigma_\eta^2)$, and $f(X_{i1t}+1, t) = F(X_{i1t}+1, t) - F(X_{i1t}, t)$.

We assume that the value functions V_{i1t} and V_{i2t} , as well as the value function differences Z_t that determine choices, are unobserved by the econometrician. The econometrician only observes the agents' choices d_{i1t} and d_{i2t} for $t=1, T$, and the wage for the chosen alternative (i.e., W_{i1t} is observed if and only if $d_{i1t}=1$). Thus, payoffs are never completely observed, both because wages are censored and because the nonpecuniary components of the payoffs ν_{i1t} and ν_{i2t} are never observed. Nevertheless, given observed choices and partially observed wages, along with the functional form assumptions about the payoff functions, it is possible to learn something about the $F(X_1, t)$ function and the structural parameters of the payoff functions without making strong assumptions about how agents form expectations.

Rather than making strong assumptions about how agents form the $F(X_1, t)$ function, we will simply allow it to be a polynomial function of the state variables X_1 and t . For example, a 3rd order polynomial would be:

$$\begin{aligned}
 (11) \quad F(X_1, t) &= \pi_0 + \pi_1 X_1 + \pi_2 X_1^2 + \pi_3 X_1^3 + \pi_4 X_1 t + \pi_5 X_1^2 t + \pi_6 X_1 t^2 \\
 &\quad + \pi_7 t + \pi_8 t^2 + \pi_9 t^3.
 \end{aligned}$$

In this case we would have:

$$\begin{aligned}
 (12) \quad f(X_{i1t}+1, t) &= \pi_1 + \pi_2(2X_{i1t}+1) + \pi_3(3X_{i1t}^2+3X_{i1t}+1) + \pi_4 t \\
 &\quad + \pi_5(2X_{i1t}+1)t + \pi_6 t^2
 \end{aligned}$$

Notice that the terms involving the parameters π_0 , π_7 , π_8 and π_9 drop out of the $f(X_{i1t}+1, t)$

function after differencing. This is because these terms capture influences on the level of the $F(X_{1,t})$ function that do not depend on the level of $X_{1,t}$. Since these terms are the same regardless of whether alternative 1 or 2 is chosen, they drop out when the future component is differenced. Also notice that the number of regressors is $P(P+1)/2$, where P is the order of the polynomial. With a 4th order polynomial, additional terms would be:

$$(4X_{i,t}^3 + 6X_{i,t}^2 + 4X_{i,t} + 1), (3X_{i,t}^2 + 3X_{i,t} + 1)t, (2X_{i,t} + 1)t^2, t^3$$

while a 5th order polynomial would add the terms:

$$(5X_{i,t}^4 + 10X_{i,t}^3 + 10X_{i,t}^2 + 5X_{i,t} + 1), (4X_{i,t}^3 + 6X_{i,t}^2 + 4X_{i,t} + 1)t, (3X_{i,t}^2 + 3X_{i,t} + 1)t^2, (2X_{i,t} + 1)t^3, t^4$$

and a 6th order polynomial adds the terms:

$$(6X_{i,t}^5 + 15X_{i,t}^4 + 20X_{i,t}^3 + 15X_{i,t}^2 + 6X_{i,t} + 1), (5X_{i,t}^4 + 10X_{i,t}^3 + 10X_{i,t}^2 + 5X_{i,t} + 1)t, (4X_{i,t}^3 + 6X_{i,t}^2 + 4X_{i,t} + 1)t^2, (3X_{i,t}^2 + 3X_{i,t} + 1)t^3, (2X_{i,t} + 1)t^4, t^5$$

In our Monte-Carlo results reported below, we use 6th order polynomials.

Writing the model in matrix notation, we have:

$$(13) \quad \ln W_{ijt} = X_{ijt}\beta_j + \epsilon_{ijt} \quad j=1,\dots,2 \quad t=1,\dots,T$$

and:

$$(14) \quad f(X_{i,t}+1,t) = S_{i,t}\pi \quad t=1,\dots,T$$

where $S_{i,t}$ is the row vector of functions of state variables that appear in the equation for $f(X_{i,t}+1,t)$ and π is the corresponding column vector of coefficients. Equation (14) highlights the fact that the polynomial function $f(X_{i,t}+1,t)$ is linear in the unknown parameters π .

The first step in a Bayesian analysis of this model via a Gibbs sampling-data augmentation algorithm is to form the "complete data" likelihood function. That is, we consider the likelihood function that could be formed if we had data on N individuals observed over T periods each, and we observed the value function differences $Z = \{Z_{it}, i=1,N; t=1,T\}$ and the

complete set of wages $W = \{(W_{i1t}, W_{i2t}), i=1, N; t=1, T\}$ for all alternatives. This is:

$$\begin{aligned}
 & \mathcal{L}(W, Z | \beta_1, \beta_2, \Sigma_\epsilon, \sigma_\eta, \pi) \\
 (15) \quad & = \prod_i \prod_t (2\pi)^{-1} |\Sigma_\epsilon|^{-1} W_{i1t}^{-1} W_{i2t}^{-1} \exp \left\{ -\frac{1}{2} \begin{bmatrix} \ell_{i1} W_{i1t} - X_{i1t} \beta_1 \\ \ell_{i2} W_{i2t} - X_{i2t} \beta_2 \end{bmatrix}' \Sigma_\epsilon^{-1} \begin{bmatrix} \ell_{i1} W_{i1t} - X_{i1t} \beta_1 \\ \ell_{i2} W_{i2t} - X_{i2t} \beta_2 \end{bmatrix} \right\} \\
 & \cdot (2\pi)^{-1/2} \sigma_\eta^{-1} \exp \left\{ -\frac{1}{2} (Z_{it} - W_{i1t} + W_{i2t} - S_{i1t} \pi)^2 / \sigma_\eta^2 \right\} \\
 & \cdot I[Z_{it} \geq 0 \text{ if } d_{i1t} = 1, \quad Z_{it} < 0 \text{ if } d_{i1t} = 0]
 \end{aligned}$$

Given flat priors, the joint posterior density of the structural parameters $(\beta_1, \beta_2, \Sigma_\epsilon, \sigma_\eta)$, the parameters characterizing expectations π , the unobserved wages $\{W_{ijt}$ when $d_{ijt}=0$ for $i=1, N, j=1, 2, t=1, T\}$, and the value function differences $\{Z_{it}, i=1, N; t=1, T\}$, is proportional to the likelihood in (15).

It is not feasible to construct this posterior density analytically, because of the high dimensional integrations over the unobserved wages and value function differences that are involved. For instance, to form the joint posterior of the parameters of interest $(\beta_1, \beta_2, \Sigma_\epsilon, \sigma_\eta, \pi)$, we must perform an $N \cdot T \cdot 2$ dimensional integration over the unobserved wages and value function differences.

But it is possible to simulate draws from the posterior using a Gibbs sampling-data augmentation algorithm. To implement this algorithm, we factor the joint posterior density (15) into a set of conditional densities, each of which can be drawn from easily. Then, we repeatedly cycle through these conditionals, drawing from each one in turn. As the number of cycles grows large, the draws so obtained converge in distribution to that of the complete joint posterior (See Gelfand and Smith (1990) or Geweke (1995b) for a discussion of convergence conditions).

Our Gibbs sampling-data augmentation algorithm consists of six steps or "blocks." To start the algorithm off, we need an initial guess for the model parameters $(\beta_1, \beta_2, \Sigma_\epsilon, \sigma_\eta, \pi)$, and for the unobserved wages. Then, the six steps are:

- (S1) Draw the value function differences $\{Z_{it}, i=1, N; t=1, T\}$
- (S2) Draw the unobserved wages $\{W_{ijt}$ when $d_{ijt}=0$ for $i=1, N, j=1, 2, t=1, T\}$
- (S3) Draw (β_1, β_2)

(S4) Draw Σ_ϵ

(S5) Draw π

(S6) Draw σ_η

Return to step S1.

Each loop through steps S1 through S6 is referred to as a "cycle" of the Gibbs sampler. We now describe the steps in detail.

Step 1

With everything else known, the Z_{it} have a simple truncated normal distribution. We have:

$$Z_{it} \sim T^+N(-W_{it} + W_{2t} - S_{it}\pi, \sigma_\eta^2) \text{ if } d_{it}=1$$

$$Z_{it} \sim T^-N(-W_{it} + W_{2t} - S_{it}\pi, \sigma_\eta^2) \text{ if } d_{it}=0$$

where T^+ indicates the normal is truncated from below at 0, while T^- indicates the normal is truncated from above at 0. It is straightforward to draw from these distributions using an inverse CDF method.

Step 2

With everything else known, the unobserved wages still have a rather complex distribution. Suppose W_{it} is unobserved. The kernel of its density is:

$$g(W_{it}) = W_{it}^{-1} \exp\{-\frac{1}{2}(\ln W_{it} - X_{it}\beta_1 - \lambda_{it})^2 / \sigma_w^2\} \exp\{-\frac{1}{2}(Z_{it} - W_{it} + W_{2t} - S_{it}\pi)^2 / \sigma_\eta^2\}$$

where $\lambda_{it} \equiv \sigma_{e_1 e_2} / \sigma_{e_2}^2$ and $\sigma_w^2 \equiv \sigma_{e_1}^2 (1 - (\sigma_{e_1 e_2} / \sigma_{e_1} \sigma_{e_2})^2)$. Notice that the appearance of log wages on the left hand side of the wage functions and the level of wages in the payoff functions creates a density that is difficult to draw from (because the inverse CDF must be constructed using quadrature methods).

Rather than using an inverse CDF method, we draw from this density using a more efficient acceptance/rejection (A/R) method described in Geweke (1995a). In this case, $g(W_{it})$

is the "target" density kernel. The "sampling" density kernel is:

$$h(W_{it}) = W_{it}^{-1} \exp\{-\frac{1}{2}(\ln W_{it} - X_{it}\beta_1 - \lambda_{it})^2/\sigma_w^2\}.$$

Note that one can draw wages from this density simply by drawing log wages from the normal density kernel $\exp\{-\frac{1}{2}(\ln W_{it} - X_{it}\beta_1 - \lambda_{it})^2/\sigma_w^2\}$ and then taking the anti-log. The acceptance probability in the A/R algorithm is:

$$\left\{ \max_w g(W)/h(W) \right\}^{-1} g(W_{it})/h(W_{it}).$$

Note that the first term in the acceptance probability is a normalizing factor that sets the maximum acceptance probability, which occurs at the point where the $g(W)/h(W)$ ratio is greatest, equal to one.

To examine the acceptance probabilities more closely in our case, define $\mu_{it} = Z_{it} + W_{it} - S_{it}\pi$. Then the ratio of the target to the sampling density kernel is:

$$g(W_{it})/h(W_{it}) = \exp\{-\frac{1}{2}(W_{it} - \mu_{it})^2/\sigma_w^2\}$$

If $\mu_{it} > 0$ then this is maximized at $W_{it} = \mu_{it}$, the normalizing factor $\left\{ \max_w g(W)/h(W) \right\}^{-1}$ is simply 1 (since the sampling density kernel is always at least as great as the target density kernel), and the acceptance probability is $g(W_{it})/h(W_{it})$. If, on the other hand, $\mu_{it} \leq 0$ then, given the constraint that wages must be positive, the ratio $g(W)/h(W)$ approaches its maximum $\exp\{-\mu_{it}^2/2\sigma_w^2\}$ as $W \rightarrow 0$. Thus, the acceptance probability is $\exp\{\mu_{it}^2/2\sigma_w^2\} g(W_{it})/h(W_{it})$. Finally, note that $\mu_{it} = W_{it} + \eta_{it}$, so that negative values will be a rather rare event provided that σ_w^2 is small relative to $\sigma_{\epsilon_1}^2$.

Step 3

With everything else known, the density of (β_1, β_2) is just:

$$\prod_i \prod_t (2\Pi)^{-1} |\Sigma_{\epsilon}|^{-1} \exp\left\{ -\frac{1}{2} \begin{bmatrix} \ell n W_{it} - X_{it} \beta_1 \\ \ell n W_{it} - X_{it} \beta_1 \\ \ell n W_{it} - X_{it} \beta_1 \\ \ell n W_{it} - X_{it} \beta_1 \end{bmatrix}' \Sigma_{\epsilon}^{-1} \begin{bmatrix} \ell n W_{it} - X_{it} \beta_1 \\ \ell n W_{it} - X_{it} \beta_1 \\ \ell n W_{it} - X_{it} \beta_1 \\ \ell n W_{it} - X_{it} \beta_1 \end{bmatrix} \right\}$$

Note that, as a function of (β_1, β_2) , this is just the kernel of a multivariate normal density. This can be constructed by estimating a seemingly unrelated regression (SUR) system in which the first equation is the $\ln W_{it}$ equation and the second is the $\ln W_{it}$ equation. Since Σ_{ϵ} is known

we apply a GLS transformation. Then (β_1, β_2) is distributed according to the standard GLS formula:

$$\beta \sim N[(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\ln W, (X'\Sigma^{-1}X)^{-1}]$$

where $\beta \equiv (\beta_1, \beta_2)'$, $\Sigma \equiv \Sigma_\epsilon \otimes I_{NT}$, X is defined as a matrix with the X_{1it} in the upper left block, X_{2it} in the lower right block, and zeros in the upper right and lower left blocks, and $\ln W$ is defined as a vector with the $\ln W_{1it}$ stacked on top and the $\ln W_{2it}$ on the bottom. It is straightforward to draw (β_1, β_2) from this bivariate normal density.

Step 4

With everything else known, Σ_ϵ^{-1} has a Wishart distribution. Define $e_{ijt} = \ln W_{ijt} - X_{ijt}\beta_j$. Then, Σ_ϵ^{-1} has the distribution:

$$\Sigma_\epsilon^{-1} \sim W[\sum_i \sum_t (e_{1it} \ e_{2it})'(e_{1it} \ e_{2it}), NT]$$

It is straightforward to draw from the bivariate Wishart, and then invert the draw for Σ_ϵ^{-1} to obtain a draw for Σ_ϵ .

Step 5

With everything else known, π has the density:

$$\prod_i \prod_t (2\pi)^{-1/2} \sigma_\eta^{-1} \exp\{-\frac{1}{2}(Z_{it} - W_{1it} + W_{2it} - S_{1it}\pi)/\sigma_\eta^2\}$$

Note that, as a function of π , this is the kernel of a multivariate normal density. This can be constructed by running a regression with the $Z_{it} - W_{1it} + W_{2it}$ as the dependent variable and the S_{1it} as the independent variables. Then π is distributed according to the standard OLS formula:

$$\pi \sim N((S'S)^{-1}S'Z^*, \sigma_\eta^2(S'S)^{-1})$$

where S is the matrix consisting of the stacked S_{1it} vectors and Z^* is the column vector consisting of the stacked $Z_{it} - W_{1it} + W_{2it}$ values. It is straightforward to draw from this using a multivariate normal random number generator.

Step 6

With everything else known, σ_η has an inverted gamma distribution. Since $\eta_{it} = \mathbf{Z}_{it}' \mathbf{W}_{it} + \mathbf{W}_{it}' \mathbf{S}_{it} \boldsymbol{\pi}$, we have that:

$$\sum_I \sum_t \eta_{it}^2 / \sigma_\eta^2 \sim \chi^2(\text{NT})$$

To draw σ_η , just draw a χ^2 random variable with NT degrees of freedom, set this equal to the quantity $\sum_I \sum_t \eta_{it}^2 / \sigma_\eta^2$, and choose σ_η accordingly.

III. Experimental Design

In order to evaluate the performance of the method described in sections I and II, we have performed several Monte-Carlo experiments. In these experiments, we generated data from the occupation choice model in equations (4)-(6), using the following parameter values:

$$\ln W_{i1t} = 9.425 + .033X_{i1t} - 0.0005X_{i1t}^2 + .000X_{i2t} + .0000X_{i2t}^2 + \epsilon_{i1t}$$

$$\ln W_{i2t} = 9.000 + .067X_{i2t} - 0.0010X_{i2t}^2 + .022X_{i1t} - .0005X_{i1t}^2 + \epsilon_{i2t}$$

$$\begin{pmatrix} \epsilon_{i1t} \\ \epsilon_{i2t} \end{pmatrix} \sim \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.0400 & 0.0000 \\ 0.0000 & 0.0625 \end{pmatrix}$$

$$\sigma_{\eta} = 400$$

We set N , the number of individuals, equal to 500, and T , the number of time periods, equal to 40.

We generated data from this model using two different assumptions about how agents form their expectations. In one case, we assume that agents solve a dynamic programming (DP) problem to generate the optimal decision rule, using a discount factor of $\delta = 0.95$. Since there are only two alternatives, and given the simplicity of the state space, it is straightforward to solve this DP problem analytically. The DP solution gives us the future components of the value functions associated with each alternative in each state in each time period. That is, we obtain:

$$F(X_{i1t} + d_{i1t}, t) = \delta E_t[V(S_i(t+1)) | S_i(t), d_{i1t}]$$

for $X_{i1t} = 0, t-1$, $d_{i1t} = 0, 1$, and $t=1, T-1$. Given these objects, we can form the optimal decision rule (10). Then, by drawing values of the stochastic terms $\{\epsilon_{i1t}, \epsilon_{i2t}, \eta_{it}\}, t=1, T, i=1, N$, we simulate lifetime occupational choice and wage paths.

A problem with using data generated in this way is that in our Gibbs sampling algorithm we constrain the $f(X_{i1t} + 1, t)$ functions to lie in a space spanned by finite order polynomials. But if the agents in the artificial data are solving a DP problem to form optimal decision rules, then neither the future components of the value functions $F(X_{i1t} + d_{i1t}, t)$, nor their differences across choices, $f(X_{i1t} + 1, t)$, will lie in a space spanned by finite order polynomials. Thus, we have a misspecification of the true data generating process.

It is of interest to see how the algorithm performs when the data generating process is correctly specified. Thus, we also generated data under the assumption that agents use a decision rule in which the future component is a polynomial in the state variables. To construct this polynomial, we regressed the $f(X_{i,t+1},t)$ functions obtained from solving the DP problem on a 6th order polynomial in the state variables. The form of the polynomial is given in equation (12) and the subsequent discussion. We then used the fitted values from this regression as the $f(X_{i,t+1},t)$ functions in the decision rule (10).

We find that the 6th order OLS fitted polynomial approximation to the future components gives a very good approximation to the optimal decision rule in two senses. First, the R^2 from the regression is .932. Second, and more importantly, we determined that agents wealth losses from using the suboptimal decision rule based on the polynomial future component rather than the optimal rule obtain by solution of the DP problem are trivial.

To determine the wealth loss from using the suboptimal decision rule we constructed two types of artificial data sets, each with $N=500$ and $T=40$. In one, agents used the optimal decision rule while in the other they used a decision rule based on the 6th order polynomial approximation to the $f(X_{i,t+1},t)$ functions. The results are reported in table 1. The data sets labeled 1 through 5 were constructed using five independent sets of draws for the ϵ_{1t} , ϵ_{2t} and η_{it} . But for each set of draws we constructed two artificial data sets. In one agents use the optimal decision rule based on the future components obtained from solution of the DP problem. Henceforth we call these data sets 1-EMAX through 5-EMAX. In the other agents use the polynomial approximation to the future components to form a decision rule. Henceforth we call these data sets 1-POLY through 5-POLY. Any differences in choices between the corresponding EMAX and POLY data sets results only because the future components of the value functions differ. Across the five experiments, the average wealth loss among the set of agents using the suboptimal rule is 8.8 hundredths of 1 percent. The finding that wealth losses from using a simple polynomial approximation to the optimal decision rule are small is consistent with the findings of Krusell and Smith (1995).

Table 2 reports on various characteristics of the artificial data sets. For data sets 1-EMAX and 1-POLY we report the fraction of agents who choose alternative 1 at each age, and the mean accepted wage in occupations 1 and 2 at each age. With these parameter values, occupation 1 can be thought of as "unskilled labor" while occupation 2 is "skilled labor." The mean of the offer wage distribution for inexperienced workers is higher in occupation 1, but in occupation 2 wages rise more quickly with experience. It is also the case that experience in occupation 1 raises offer wages in occupation 2, but not vice-versa. Together, these features

create an "occupational ladder" in which workers have some tendency to shift from 1 to 2 as they get older.

Table 3 reports the results from applying the Gibbs sampling algorithm described in section II to the POLY data sets. After a sufficient number of "burn-in" cycles so as to achieve convergence (with the criteria to be discussed below), we used the next 2000 cycles to simulate the joint posterior distribution. The column labeled "True" in table 3 contains the data generating values for the structural parameters $(\beta_1, \beta_2, \Sigma_\epsilon, \sigma_\eta)$ and the 21 polynomial coefficients π that capture expectations. The column labeled "Mean" contains the posterior means, followed in the subsequent columns by the posterior median, posterior standard deviation, and the minimum and maximum draw.

Overall, the results in table 3 are quite impressive. For instance, in the runs on the 1-POLY data, 2-POLY and 4-POLY data sets, the posterior means for all the wage equation parameters are all within two posterior standard deviations of the data generating values, and most are well within one standard deviation. The posterior mean for the wage error correlation is slightly more than two posterior standard deviations above the data generating value in the run on the 3-POLY data set, while the posterior mean for the occupation 2 wage error variance is slightly more than two posterior standard deviations above the data generating value in the run on the 5-POLY data set.

Table 4 contains OLS regressions of accepted log wages on experience, ignoring the dynamic selection bias that is generated by agents' decision rule, for the 1-POLY through 5-POLY data sets. These estimates show substantial biases for all the wage equation parameters. Thus, it is clear that the Gibbs sampling algorithm is doing an impressive job of implementing the appropriate dynamic selection correction.

The results in table 3 also indicate that there is difficulty in pinning down σ_η , the standard deviation of the nonpecuniary component of payoffs. In the runs on data sets 1-POLY, 2-POLY and 4-POLY, the posterior means for σ_η are well above the data generating value. However, some investigation revealed that choices are quite insensitive to rather large changes in σ_η . Recall that the true value of σ_η is 400. In an experiment we regenerated one of the artificial data sets using the same sets of draws for the ϵ_{11t} , ϵ_{22t} and η_{it} , but scaling σ_η up to 600. In that case, 97.4 percent of choices remained identical. When we scaled σ_η down to 200, 97.4 percent of choices also remained identical. Thus, it is not surprising that σ_η is hard to identify.

Finally, we note that for the most part the algorithm does a good job of uncovering the true values of the polynomial coefficients π that capture expectations. But a close examination reveals that we tend to better on the lower order polynomial terms. In particular, the draws for

π_{21} , the coefficient on the 6th order experience term, are usually too high. This suggests we may have some trouble capturing the future components that agents assign to choices near the end of the life cycle. (Some evidence to this effect is reported below).

More interesting is an analysis of how the algorithm performs in the case where agents do in fact use the optimal decision rule. Table 5 reports the results for artificial data sets 1-EMAX through 5-EMAX. The column labeled "True" still contains the data generating values of the structural parameters. However, for the π coefficients we simply reproduce the coefficients from the OLS regression of the $f(X_{i,t}+1,t)$ functions obtained by solving the DP problem on a 6th order polynomial in the state variables. Thus, these π values are only reported so that we can gauge the extent to which the Gibbs sampling algorithm obtains a polynomial fit to the future components that looks similar to the OLS fit.

The results in table 5 are again quite impressive. There are a few instances in which the structural parameters in the wage equations are (slightly) more than two posterior standard deviations away from the data generating values. These include: 1) the occupation 2 own-experience linear and quadratic terms in data set 1-EMAX, 2) the occupation 2 cross-experience linear and quadratic terms in data sets 1-EMAX and 5-EMAX, 3) and the occupation 1 cross-experience linear and quadratic terms in data set 3-EMAX. But these four instances are not of much concern, because in each case where the linear term in the quadratic function is too big (small), the quadratic term is also too big (small). Such problems in pinning down the exact curvature of quadratic functions is common in data sets of this size, even when the data generating process is correctly specified. The posterior mean for the wage error correlation is also slightly more than two posterior standard deviations below the data generating value in data sets 1-EMAX and 4-EMAX.

Table 6 contains OLS regressions of accepted log wages on experience, ignoring the dynamic selection bias that is generated by agents' decision rule, for the 1-EMAX through 5-EMAX data sets. These estimates again show substantial biases for all the wage equation parameters. Thus, the Gibbs sampling algorithm continues to do an impressive job of implementing a dynamic selection correction despite the fact that agents' decision rule is misspecified due to the 6th order polynomial approximation.

Returning to table 5, we see that we again have difficulty in pinning down σ_η , just as with the POLY data sets. Its posterior mean is well above the data generating value in the runs on data sets 2-EMAX, 4-EMAX and 5-EMAX. But again, it appears that choices are little affected by changes in the magnitude of σ_η on the order we observe here.

Figure 1 contains the simulated posterior densities for each of the structural parameters,

using data set 3-EMAX. In each figure, a vertical bar indicates the true data generating value of the parameter. Tick marks along the horizontal axis indicate the posterior mean and points one and two posterior standard deviations away from the mean. Except for the cross-experience and cross-experience squared coefficients in occupation 1, the posterior means are all very close to the data generating values. An interesting feature is the bimodal density for σ_{η} .

Figure 2A reports the joint posterior density for the linear and quadratic terms in occupation 1 own experience for data set 3-EMAX, while figure 2B reports the same information for the cross experience terms. These figures illustrate the strong negative correlation between the linear and quadratic terms, which is the source of the problem in pinning down these terms individually that was mentioned above.

Returning to table 5, a comparison of the posterior means of the π parameters with the OLS estimates obtained by regression of the actual $f(X_{it}+1,t)$ functions on a 6th order polynomial in the state variables reveals many substantial differences. Thus, the polynomial approximation to the future components in the agents' optimal decision rule generated by the Gibbs algorithm appears (superficially) to be quite different from that generated by OLS regression.

Table 7 contains an analysis of how well the approximation to agents' decision rule obtained by the Gibbs algorithm matches the agents' optimal decision rule. The following experiment was performed: For each of the data sets 1-EMAX through 5-EMAX, we formed the polynomial approximation to the $f(X_{it}+1,t)$ function implied by the posterior means for the π parameters obtained from the respective runs. We then used these polynomial approximations to the $f(X_{it}+1,t)$ functions to form an approximate versions of agents' decision rule (10). We then regenerated each of the 5 data sets, using the exact same draws for the ϵ_{it} , ϵ_{2t} and η_{it} , but with the polynomial approximations replacing the optimal $f(X_{it}+1,t)$ functions in (10). The average wealth loss across the 5 data sets if agents use these suboptimal decision rules is 7.6 hundredths of 1 percent. This is actually slightly smaller than the average wealth loss of 8.8 hundredths of 1 percent obtained when OLS estimates of the π parameters were used (see table 1).

Figures 3, 4 and 5 compare the true values and polynomial approximations to the $f(X_{it}+1,t)$ - the difference between the future component associated with choice of occupation 1 and that associated with choice of 2. Figure 3A plots the true values of the $f(X_{it}+1,t)$ at various state points that arise in periods 1 through 5, while figure 3B plots the polynomial approximation. The horizontal axis shows the level of occupation 1 experience at each state point.

For example, in period 1 the only possible state is zero experience in occupation 1. That point is indicated by a box in figures 3A and 3B. In that state, the true value of $f(X_{it}+1,t) = f(1,1)$ is minus \$2653. The value is negative because the return to experience is less in occupation 1 than in occupation 2. The polynomial approximation is minus \$2523.

In period 2 there are two possible states - either one or zero periods of experience in occupation 1. These two points are indicated by shaded hour glasses in figures 3A and 3B. When occupation 1 experience is zero, the true value of $f(X_{it}+1,t) = f(1,2)$ is minus \$3022. The polynomial approximation is minus \$2945. When occupation 1 experience is one, the true value of $f(X_{it}+1,t) = f(2,2)$ is minus \$2805. The polynomial approximation is minus \$2837.

Overall, figures 3, 4 and 5 show that the Gibbs sampling algorithm does an excellent job of uncovering agents' expectations. As figure 4C indicates, the polynomial approximations to the $f(X_{it}+1,t)$ functions are generally within a few hundred dollars of the true values at state points that may arise over the first periods. Figure 4C also shows that the accuracy of the approximation deteriorates over certain regions of the state space that may arise in later periods.

This pattern, whereby the inferences about the future components of the value functions deteriorate with age, is to be expected for two reasons: First, since in later periods there are more possible states, the data will tend to contain fewer agents at each state point. Thus, there is less data to pin down the $f(X_{it}+1,t)$ functions at those less frequently observed state points. Second, as agents age they tend to specialize in either occupation 1 or 2. As a result of human capital accumulation, the current payoff function realizations for that occupation in which an agent specializes come more and more to dominate those for the alternative occupation. Thus, with age, choices become less influenced by the future components, and the data therefore contains less information about them.

Despite the fact that our inferences about the future components of the value functions tend to deteriorate with age (at least over certain regions of the state space), figures 5A and 5B indicate that the Gibbs sampling algorithm does a good job of picking up the overall shape of the $f(X_{it}+1,t)$ functions over the whole life cycle.

Finally, figure 6 contains graphical evidence on convergence of the Gibbs sampling algorithm for data set 3-EMAX. To gauge convergence we used such graphical evidence. Specifically, we constructed charts of the draws by cycle and the posterior means by cycle, and let the Gibbs sampler run until it appeared that the distribution was stationary. We found that from 5,000 to 18,000 cycles were necessary to achieve convergence, with substantial variation across runs. For data set 3-EMAX, we concluded that convergence had been achieved after 18,000 cycles. Each cycle required approximately 12.6 seconds on an IBM RISC 6000 model

360. That is 0.0252 seconds per person, and 0.00063 seconds per person-year.

V. Conclusion

In this paper we have illustrated a method for Bayesian inference in dynamic discrete choice models that does not require the econometrician to make strong assumptions about how agents form expectations, or to solve the agents' dynamic optimization problem. In this method, we use observed data on agents' choices and payoffs to infer both the structural parameters of the payoff functions that agents confront, and the agents' expectations of the values of occupying different states. The results of our numerical application to artificial data generated from a human capital based occupational choice model indicate that the method does a good job both of drawing inferences about the structural parameters of the payoff functions and of drawing inferences about expectations.

Our results indicate that misspecification of agents' decision rules, in the form of assuming that the future components of the value functions follow a 6th order polynomial when in fact they are the discounted values of the EMAX functions that come from a solution of a dynamic programming problem, is a "second order" problem in two senses. First, it has a negligible effect on inferences with regard to the structural parameters of the payoff functions. Second, the misspecified decision rule that we infer from the data is very close to the optimal rule - in the sense that if agents were to use our suboptimal rule rather than the optimal rule it would lead to trivial wealth losses.

In future work, we plan to investigate how our method performs if the current payoff functions are misspecified. Then, we will compare this to the performance of full solution methods under misspecification of the current payoff functions. We conjecture that in statistical inference for dynamic optimization models the misspecification of the current payoff functions is a "first order" problem relative to the misspecification of the future component of the value functions.

The Gibbs sampling-data augmentation approach to Bayesian inference that we have used here has previously been successfully applied to inference in static discrete choice models in work by McCulloch and Rossi (1994) and Geweke, Keane and Runkle (1994a,b), among others. The results here indicate that the method also works well in selection type models. This outcome was not obvious from the previous results because of potential problems associated with the step of the Gibbs algorithm in which the unobserved wages are drawn. As we have described, this step is significantly more complex than the step in which latent utilities are drawn in a pure discrete choice framework without observed payoffs.

References

- Bellman, R. (1957). *Dynamic Programming*, Princeton: Princeton University Press.
- Gelfand, A.E., and A.F.M. Smith (1990). "Sampling Based Approaches to Calculating Marginal Densities," Journal of the American Statistical Association, 85: 398-409.
- Geweke, J. (1995a). "Priors for Macroeconomic Time Series and Their Application." Working paper, Federal Reserve Bank of Minneapolis.
- Geweke, J. (1995b). "Simulation Based Bayesian Inference for Economic Time Series." Working paper, Federal Reserve Bank of Minneapolis.
- Geweke, J., Keane, M. and D. Runkle (1994a). "Alternative Computational Approaches to Inference in the Multinomial Probit Model," Review of Economics and Statistics, 609-632.
- Geweke, J., Keane, M. and D. Runkle (1994b). "Statistical Inference in the Multinomial Probit Model," Staff Report 177, Federal Reserve Bank of Minneapolis.
- Hotz, V.J. and R.A. Miller (1993). "Conditional Choice Probabilities and the Estimation of Dynamic programming Models," Review of Economic Studies, 60:497-530.
- Keane, M. and K. Wolpin (1994). "Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," Review of Economics and Statistics, 648-672.
- Krusell, P. and A.A. Smith Jr. (1995). "Rules of Thumb in Macroeconomic Equilibrium: A Quantitative Analysis," Journal of Economic Dynamics and Control, forthcoming.
- Lancaster, T. (1995) "Exact Structural Inference in Optimal Job Search Models," Forthcoming, Journal of Business and Economic Statistics.
- Manski, C. (1993). "Dynamic Choice in Social Settings," Journal of Econometrics, 58: 121-136.
- McCulloch, R., and P. Rossi (1994). "An Exact Likelihood Analysis of the Multinomial Probit Model," Journal of Econometrics, 64:1, 207-240.
- Rust, J. (1995) "Structural Estimation of Markov Decision Processes," in R. Engle and D. McFadden (eds.) Handbook of Econometrics, Vol. 4 (Amsterdam: North Holland).
- Tanner, M.A. and W.H. Wong (1987) "The Calculation of Posterior Distributions by Data Augmentation," Journal of the American Statistical Association, 82: 528-550.

Table 1

Quality of the OLS Estimated Polynomial Approximation to
the True Future Component

Error Set	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
Mean* Wealth** with True Future Component	354534.72	354912.59	356430.12	355723.35	355438.58
Mean* Wealth** with Polynomial Approximation	354404.31	354401.22	356092.88	355612.60	354926.21
Mean* Dollar Loss	130.43	511.37	337.24	110.75	512.37
Mean* Percent Loss	0.03%	0.14%	0.09%	0.03%	0.14%
Percent Agreement in Choices					
Aggregate	92.40%	92.80%	90.60%	93.20%	92.80%
By Age					
1	97.60%	97.80%	95.00%	96.20%	96.80%
2	98.00%	98.80%	97.40%	97.20%	98.20%
3	98.60%	98.40%	97.60%	98.60%	98.20%
4	96.60%	96.60%	95.40%	97.40%	97.20%
5	94.40%	95.20%	93.20%	96.00%	97.00%
10	90.40%	88.60%	89.80%	91.40%	87.20%
20	91.20%	93.00%	91.20%	92.60%	92.40%
30	91.80%	92.20%	91.40%	94.60%	93.60%
40	92.60%	93.40%	90.60%	93.20%	92.80%

* Mean is taken over 500 agents that live for exactly 40 periods.

** "Wealth" is defined as the discounted stream of ex-post lifetime earnings.

Table 2

Choice Distributions and Mean Accepted Wages in the Data
Generated with True and OLS Polynomial Future Components

Period	<u>Data Set 1-EMAX</u>			<u>Data Set 1-POLY</u>		
	<u>Mean Accepted Wage</u>			<u>Mean Accepted Wage</u>		
	<u>Percent in Occ. 1</u>	<u>Occ. 1</u>	<u>Occ. 2</u>	<u>Percent in Occ. 1</u>	<u>Occ. 1</u>	<u>Occ. 2</u>
1	68.20%	13518.95	10157.72	65.80%	13582.70	10090.85
2	64.20%	14075.90	10333.34	62.20%	14143.49	10270.84
3	63.60%	14259.73	10412.94	62.60%	14273.48	10410.98
4	61.60%	14711.21	10720.55	60.20%	14722.20	10736.39
5	57.80%	15149.64	11265.89	56.60%	15126.34	11352.12
6	53.40%	15074.74	11386.17	52.40%	15024.08	11511.15
7	51.20%	15873.34	12043.19	49.80%	15916.81	12104.61
8	53.60%	16110.32	12483.63	50.60%	16366.56	12351.35
9	46.80%	16435.97	12807.57	44.40%	16580.96	12849.18
10	45.40%	16880.64	13191.99	43.40%	17012.38	13276.41
11	43.20%	17147.20	13546.29	40.20%	17178.43	13654.06
12	44.80%	17228.21	14193.09	40.40%	17349.29	14210.80
13	37.80%	17581.22	15073.76	36.20%	17702.36	15086.65
14	37.80%	17434.20	15696.70	34.40%	17389.17	15704.33
15	36.80%	17880.31	16124.57	35.20%	17708.52	16266.96
16	32.60%	18237.00	16864.88	31.60%	18031.73	16956.84
17	31.80%	18776.79	17215.52	29.60%	18524.06	17288.61
18	33.00%	18915.48	18427.82	31.80%	18754.68	18460.48
19	32.40%	18929.04	18692.19	30.40%	19070.55	18728.53
20	26.60%	19465.56	19474.62	26.20%	19173.03	19581.50
21	30.20%	20174.20	20024.03	29.60%	20105.41	20154.48
22	30.80%	19771.71	20891.76	29.80%	19752.06	21014.38
23	28.20%	19931.67	21044.38	28.20%	19888.64	21126.76
24	27.20%	19612.29	22108.70	27.00%	19896.59	22160.10
25	27.40%	20484.94	22476.02	27.60%	20275.29	22661.92
26	26.80%	19880.57	23026.35	26.00%	19946.60	22899.08
27	26.40%	21125.67	23990.42	25.60%	21112.75	23911.17
28	27.00%	20889.27	23755.55	28.20%	20743.88	23989.04
29	25.40%	20784.30	25539.36	24.60%	20844.08	25487.78
30	26.60%	21159.77	25761.60	25.20%	21157.22	25691.86
31	25.60%	21421.28	26887.25	25.60%	21302.05	26904.03
32	24.40%	21271.78	26857.47	23.80%	21276.90	26786.20
33	24.00%	20470.71	27392.54	22.80%	20497.13	27373.38
34	23.20%	22078.95	27592.74	22.20%	22101.79	27459.55
35	24.80%	21493.14	27957.00	22.80%	22002.08	27700.57
36	23.40%	21680.03	27799.76	22.20%	22047.80	27706.19
37	23.80%	21546.08	28819.30	22.20%	21896.35	28533.66
38	22.20%	21255.37	29480.34	21.00%	21475.40	29466.02
39	21.20%	20816.65	29277.87	20.80%	20892.50	29241.84
40	22.00%	21680.48	28889.35	20.60%	21976.99	28583.09

Table 3

Descriptive Statistics for Final 2000 Gibbs Sampler Parameter Draws for Several
Different Data Sets Generated Using Polynomial Future Component

Data Set 1-Polynomial

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42581	9.42574	0.00480	9.41139	9.44243
Occ. 1 Own Exp.	0.03300	0.03262	0.03260	0.00074	0.03005	0.03516
Occ. 1 Own Exp. Sq.	-0.00050	-0.00049	-0.00049	0.00002	-0.00059	-0.00043
Occ. 1 Other Exp.	0.00000	0.00102	0.00103	0.00099	-0.00197	0.00385
Occ. 1 Other Exp Sq.	0.00000	-0.00005	-0.00005	0.00003	-0.00014	0.00006
Occ. 1 Error Var.	0.04000	0.03873	0.03876	0.00084	0.03618	0.04116
Occ. 2 Int.	9.00000	9.00645	9.00668	0.00662	8.98432	9.03120
Occ. 2 Own Exp.	0.06700	0.06763	0.06764	0.00100	0.06461	0.07070
Occ. 2 Own Exp. Sq.	-0.00100	-0.00102	-0.00102	0.00003	-0.00111	-0.00093
Occ. 2 Other Exp.	0.02200	0.01960	0.01961	0.00139	0.01453	0.02323
Occ. 2 Other Exp Sq.	-0.00050	-0.00043	-0.00043	0.00005	-0.00057	-0.00026
Occ. 2 Error Var.	0.06250	0.06245	0.06244	0.00087	0.05927	0.06540
Error Covariance	0.00000	-0.00154	-0.00150	0.00119	-0.00490	0.00177
Error Correlation	0.00000	-0.03122	-0.03046	0.02400	-0.09783	0.03692
SD Eta [400.00	877.92	877.41	89.37	678.54	1045.30
PI 1	-2861.22	-2569.51	-2560.35	219.86	-3316.45	-1921.61
PI 2	-717.03	-2789.23	-2811.85	590.09	-4242.62	-1034.00
PI 3	2285.40	3883.24	3896.59	497.74	2274.22	5283.83
PI 4	-790.83	-1181.06	-1187.38	187.77	-1670.83	-539.21
PI 5	108.03	146.48	147.96	29.91	49.33	228.67
PI 6	-5.24	-6.57	-6.73	1.62	-11.01	-1.72
PI 7	137.86	1483.24	1510.31	376.34	342.40	2571.53
PI 8	-5757.05	-6719.48	-6674.38	767.26	-9364.73	-4448.95
PI 9	2835.75	2897.92	2933.42	403.07	1753.28	4081.22
PI10	-437.15	-429.25	-438.14	79.04	-636.24	-213.40
PI11	21.28	20.94	21.34	4.82	8.04	32.91
PI12	3763.28	3612.86	3627.42	403.40	2471.73	4959.95
PI13	-2986.69	-2773.78	-2787.23	386.38	-3758.32	-1896.12
PI14	479.48	485.97	478.12	92.89	279.73	716.27
PI15	-20.18	-24.70	-23.84	6.48	-43.35	-9.62
PI16	897.49	882.39	872.97	153.68	479.70	1331.18
PI17	-88.54	-161.07	-160.47	49.31	-313.72	-41.65
PI18	-4.20	6.95	7.06	4.63	-4.73	21.68
PI19	-57.82	-22.99	-23.23	15.22	-64.18	14.24
PI20	8.92	2.58	2.60	1.84	-2.66	7.87
PI21	-0.74	0.00	0.00	0.20	-0.72	0.78

Data Set 2-Polynomial

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42975	9.42956	.00472	9.41350	9.44386
Occ. 1 Own Exp.	.03300	.03182	.03184	.00076	.02913	.03444
Occ. 1 Own Exp. Sq.	-.00050	-.00047	-.00047	.00002	-.00054	-.00039
Occ. 1 Other Exp.	.00000	.00036	.00038	.00088	-.00234	.00316
Occ. 1 Other Exp Sq.	.00000	-.00001	-.00001	.00004	-.00011	.00010
Occ. 1 Error Var.	.04000	.03960	.03956	.00072	.03730	.04224
Occ. 2 Int.	9.00000	9.01192	9.01213	.00649	8.98975	9.03430
Occ. 2 Own Exp.	.06700	.06585	.06587	.00081	.06275	.06822
Occ. 2 Own Exp. Sq.	-.00100	-.00098	-.00098	.00002	-.00106	-.00089
Occ. 2 Other Exp.	.02200	.02205	.02201	.00122	.01796	.02609
Occ. 2 Other Exp Sq.	-.00050	-.00051	-.00051	.00005	-.00068	-.00035
Occ. 2 Error Var.	.06250	.06087	.06088	.00087	.05766	.06371
Error Covariance	.00000	.00073	.00068	.00085	-.00145	.00316
Error Correlation	.00000	.01494	.01410	.01721	-.02957	.06393
SD Eta [400.00	626.78	598.40	79.67	506.08	802.73
PI 1	-2861.22	-2831.80	-2840.92	187.41	-3372.70	-2237.46
PI 2	-717.03	-1238.40	-1233.04	619.12	-2893.84	589.55
PI 3	2285.40	3026.43	3013.73	504.76	1355.90	4489.56
PI 4	-790.83	-1021.39	-1012.77	152.72	-1520.02	-527.63
PI 5	108.03	127.16	126.22	19.56	67.60	194.44
PI 6	-5.24	-5.44	-5.41	.93	-8.44	-2.84
PI 7	137.86	1029.23	1047.37	397.84	-178.19	2217.27
PI 8	-5757.05	-6936.46	-6938.52	655.01	-8521.04	-5019.01
PI 9	2835.75	3098.48	3094.69	246.50	2349.86	3813.03
PI10	-437.15	-415.57	-418.96	39.49	-527.39	-277.77
PI11	21.28	17.25	17.60	2.52	9.57	24.17
PI12	3763.28	3907.54	3915.65	328.78	3006.69	4682.20
PI13	-2986.69	-2845.31	-2862.74	204.96	-3473.74	-2295.56
PI14	479.48	357.96	366.56	58.12	185.12	476.69
PI15	-20.18	-9.82	-10.19	4.88	-20.26	4.23
PI16	897.49	802.56	817.72	110.10	465.96	1054.95
PI17	-88.54	7.72	8.01	42.97	-98.82	130.25
PI18	-4.20	-11.02	-10.99	4.35	-23.64	.44
PI19	-57.82	-81.97	-80.47	14.31	-124.11	-49.66
PI20	8.92	9.22	9.09	1.67	4.88	14.80
PI21	-.74	.00	.00	.16	-.48	.57

Data Set 3-Polynomial

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42433	9.42428	0.00463	9.40993	9.43992
Occ. 1 Own Exp.	0.03300	0.03339	0.03340	0.00076	0.03085	0.03592
Occ. 1 Own Exp. Sq.	-0.00050	-0.00050	-0.00050	0.00002	-0.00058	-0.00043
Occ. 1 Other Exp.	0.00000	0.00179	0.00161	0.00105	-0.00067	0.00572
Occ. 1 Other Exp Sq.	0.00000	-0.00006	-0.00005	0.00004	-0.00020	0.00003
Occ. 1 Error Var.	0.04000	0.03923	0.03921	0.00062	0.03713	0.04122
Occ. 2 Int.	9.00000	9.00963	9.00971	0.00596	8.99033	9.02821
Occ. 2 Own Exp.	0.06700	0.06695	0.06694	0.00087	0.06439	0.06985
Occ. 2 Own Exp. Sq.	-0.00100	-0.00101	-0.00101	0.00003	-0.00109	-0.00093
Occ. 2 Other Exp.	0.02200	0.02210	0.02205	0.00135	0.01751	0.02611
Occ. 2 Other Exp Sq.	-0.00050	-0.00054	-0.00054	0.00005	-0.00069	-0.00037
Occ. 2 Error Var.	0.06250	0.06177	0.06172	0.00091	0.05870	0.06464
Error Covariance	0.00000	0.00264	0.00253	0.00073	0.00090	0.00460
Error Correlation	0.00000	0.05376	0.05148	0.01502	0.01793	0.09445
SD Eta	400.00	365.07	364.19	23.43	318.61	439.33
PI 1	-2861.22	-2770.31	-2777.93	233.39	-3483.21	-2067.96
PI 2	-717.03	-1247.45	-1292.60	634.46	-2860.29	468.68
PI 3	2285.40	2322.31	2321.32	458.09	1051.95	3507.73
PI 4	-790.83	-667.07	-660.94	159.63	-1137.73	-275.26
PI 5	108.03	79.06	77.80	23.52	26.01	145.97
PI 6	-5.24	-3.51	-3.37	1.21	-6.63	-0.94
PI 7	137.86	425.65	383.70	468.02	-772.88	1658.89
PI 8	-5757.05	-5055.43	-5098.36	682.36	-6733.75	-3301.32
PI 9	2835.75	2158.62	2092.33	306.03	1491.26	2980.31
PI10	-437.15	-305.88	-291.50	50.34	-433.52	-219.99
PI11	21.28	14.22	13.76	2.98	8.25	21.05
PI12	3763.28	3160.24	3134.61	318.48	2367.18	3868.82
PI13	-2986.69	-2267.32	-2159.70	285.97	-3054.56	-1740.23
PI14	479.48	353.01	355.79	76.52	201.23	515.07
PI15	-20.18	-15.51	-16.31	5.81	-27.39	-2.77
PI16	897.49	704.08	704.53	121.60	374.88	946.47
PI17	-88.54	-80.92	-89.28	52.39	-192.48	37.69
PI18	-4.20	0.18	1.09	4.77	-11.12	10.82
PI19	-57.82	-38.75	-38.10	12.65	-68.47	-8.70
PI20	8.92	4.37	4.31	1.44	0.87	8.12
PI21	-0.74	0.00	0.00	0.09	-0.31	0.35

Data Set 4-Polynomial

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42540	9.42547	.00499	9.40842	9.44164
Occ. 1 Own Exp.	.03300	.03293	.03292	.00078	.03040	.03536
Occ. 1 Own Exp. Sq.	-.00050	-.00050	-.00050	.00002	-.00057	-.00042
Occ. 1 Other Exp.	.00000	.00130	.00136	.00118	-.00221	.00455
Occ. 1 Other Exp Sq.	.00000	-.00007	-.00007	.00004	-.00017	.00006
Occ. 1 Error Var.	.04000	.03956	.03951	.00088	.03734	.04306
Occ. 2 Int.	9.00000	8.99729	8.99712	.00725	8.97404	9.02495
Occ. 2 Own Exp.	.06700	.06876	.06873	.00092	.06592	.07195
Occ. 2 Own Exp. Sq.	-.00100	-.00104	-.00104	.00003	-.00113	-.00095
Occ. 2 Other Exp.	.02200	.02076	.02073	.00136	.01647	.02504
Occ. 2 Other Exp Sq.	-.00050	-.00051	-.00051	.00007	-.00071	-.00029
Occ. 2 Error Var.	.06250	.06154	.06156	.00089	.05855	.06466
Error Covariance	.00000	-.00174	-.00147	.00135	-.00618	.00136
Error Correlation	.00000	-.03505	-.02989	.02682	-.12191	.02781
SD Eta [400.00	674.58	724.84	113.83	464.21	870.08
PI 1	-2861.22	-2536.40	-2550.90	198.81	-3162.43	-1869.93
PI 2	-717.03	-1899.82	-1866.77	597.81	-3981.71	-324.84
PI 3	2285.40	3274.53	3279.02	541.34	1878.76	5294.37
PI 4	-790.83	-1142.62	-1154.73	227.35	-1802.93	-509.60
PI 5	108.03	153.39	154.55	39.35	54.00	254.71
PI 6	-5.24	-7.11	-7.12	2.19	-12.19	-1.56
PI 7	137.86	278.15	287.62	419.08	-901.18	1643.12
PI 8	-5757.05	-5916.11	-5951.45	780.33	-8438.24	-3608.96
PI 9	2835.75	3034.69	3042.88	495.56	1902.69	4344.79
PI10	-437.15	-459.75	-450.83	98.94	-711.45	-243.92
PI11	21.28	21.68	21.03	5.84	8.84	36.60
PI12	3763.28	3833.76	3885.07	460.40	2655.19	5097.36
PI13	-2986.69	-3270.66	-3190.73	446.07	-4447.08	-2364.28
PI14	479.48	520.19	503.08	98.28	300.93	806.76
PI15	-20.18	-22.55	-22.16	6.26	-42.31	-8.80
PI16	897.49	1025.35	1005.58	146.49	669.54	1456.10
PI17	-88.54	-110.93	-112.27	41.10	-237.68	-7.12
PI18	-4.20	-.55	-.30	3.60	-11.21	9.76
PI19	-57.82	-56.08	-54.55	10.77	-94.00	-29.06
PI20	8.92	6.34	6.18	1.31	2.81	11.44
PI21	-.74	-.01	-.01	.16	-.58	.47

Data Set 5-Polynomial

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42477	9.42472	0.00470	9.40929	9.44211
Occ. 1 Own Exp.	0.03300	0.03296	0.03298	0.00076	0.03054	0.03549
Occ. 1 Own Exp. Sq.	-0.00050	-0.00050	-0.00050	0.00002	-0.00057	-0.00043
Occ. 1 Other Exp.	0.00000	0.00076	0.00073	0.00104	-0.00208	0.00456
Occ. 1 Other Exp Sq.	0.00000	-0.00004	-0.00004	0.00004	-0.00018	0.00007
Occ. 1 Error Var.	0.04000	0.03985	0.03986	0.00069	0.03736	0.04238
Occ. 2 Int.	9.00000	8.99958	8.99951	0.00605	8.97836	9.01842
Occ. 2 Own Exp.	0.06700	0.06761	0.06763	0.00092	0.06416	0.07111
Occ. 2 Own Exp. Sq.	-0.00100	-0.00101	-0.00102	0.00003	-0.00112	-0.00092
Occ. 2 Other Exp.	0.02200	0.02077	0.02074	0.00127	0.01648	0.02466
Occ. 2 Other Exp Sq.	-0.00050	-0.00047	-0.00047	0.00005	-0.00063	-0.00030
Occ. 2 Error Var.	0.06250	0.06464	0.06463	0.00089	0.06183	0.06767
Error Covariance	0.00000	-0.00103	-0.00104	0.00072	-0.00358	0.00093
Error Correlation	0.00000	-0.02031	-0.02052	0.01404	-0.07029	0.01823
SD Eta	400.00	594.25	557.47	128.49	406.23	815.81
PI 1	-2861.22	-2823.83	-2805.36	217.86	-3476.93	-2186.27
PI 2	-717.03	-1374.43	-1412.63	547.50	-2886.53	276.85
PI 3	2285.40	3054.51	3043.68	531.96	1513.29	4599.72
PI 4	-790.83	-1038.84	-1034.14	189.35	-1631.32	-526.30
PI 5	108.03	133.97	135.10	27.80	60.67	218.46
PI 6	-5.24	-6.00	-6.08	1.48	-10.06	-2.32
PI 7	137.86	845.36	836.20	388.94	-163.09	1832.36
PI 8	-5757.05	-6881.36	-6898.43	747.04	-8901.33	-4814.42
PI 9	2835.75	3233.48	3244.04	339.08	2276.31	4172.79
PI10	-437.15	-458.84	-458.11	64.93	-640.20	-284.09
PI11	21.28	20.34	19.74	4.37	10.49	34.79
PI12	3763.28	4110.59	4123.52	324.99	3038.34	5013.84
PI13	-2986.69	-3191.07	-3161.70	270.15	-4116.02	-2538.56
PI14	479.48	436.44	416.69	87.82	292.87	783.30
PI15	-20.18	-14.51	-13.18	7.28	-42.01	-0.27
PI16	897.49	930.51	905.96	120.27	606.35	1370.58
PI17	-88.54	-10.34	-8.62	60.63	-225.97	129.68
PI18	-4.20	-10.70	-10.60	5.70	-24.25	8.04
PI19	-57.82	-91.47	-90.78	14.01	-128.67	-50.28
PI20	8.92	10.36	10.30	1.66	5.12	15.37
PI21	-0.74	-0.02	-0.02	0.19	-0.72	0.64

Table 4

OLS Wage Equation Parameter Estimates for Data Generated with Polynomial Future Component (standard deviations in parentheses)*

Data Set	Occupation One					Occupation Two					Wage Error SDs	
	Intercept	Own Exp.	Own Exp. Squared	Other Exp.	Other Exp. Squared	Intercept	Own Exp.	Own Exp. Squared	Other Exp.	Other Exp. Squared	Occ. 1	Occ. 2
Actual	9.42500	0.03300	-0.00050	0.00000	0.00000	9.00000	0.06700	-0.00100	0.02200	-0.00050	0.20000	0.25000
1-POLY	9.51154 (0.00459)	0.02621 (0.00071)	-0.00039 (0.00002)	0.01970 (0.00104)	-0.00049 (0.00005)	9.17053 (0.00616)	0.05818 (0.00088)	-0.00083 (0.00003)	0.01124 (0.00148)	0.00012 (0.00007)	0.17728	0.22841
2-POLY	9.51300 (0.00466)	0.02633 (0.00077)	-0.00040 (0.00002)	0.01991 (0.00103)	-0.00050 (0.00004)	9.17570 (0.00620)	0.05665 (0.00087)	-0.00079 (0.00003)	0.01210 (0.00163)	0.00011 (0.00008)	0.17662	0.22648
3-POLY	9.51180 (0.00463)	0.02707 (0.00073)	-0.00041 (0.00002)	0.02029 (0.00104)	-0.00053 (0.00005)	9.16835 (0.00619)	0.05781 (0.00087)	-0.00082 (0.00003)	0.01306 (0.00157)	0.00008 (0.00008)	0.17806	0.22836
4-POLY	9.51410 (0.00460)	0.02648 (0.00072)	-0.00040 (0.00002)	0.02039 (0.00102)	-0.00051 (0.00004)	9.16211 (0.00623)	0.05899 (0.00088)	-0.00084 (0.00003)	0.01213 (0.00161)	0.00007 (0.00008)	0.17732	0.22673
5-POLY	9.51402 (0.00472)	0.02696 (0.00077)	-0.00041 (0.00002)	0.01999 (0.00101)	-0.00053 (0.00004)	9.17109 (0.00625)	0.05834 (0.00089)	-0.00082 (0.00003)	0.00881 (0.00170)	0.00031 (0.00009)	0.17676	0.23088

*Results from simple OLS regression using only observed wages and experience.

Table 5

Descriptive Statistics for Final 2000 Gibbs Sampler Parameter Draws for Several
Different Data Sets Generated Using True Future Component

Data Set 1-EMAX

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42126	9.42126	.00484	9.40353	9.43603
Occ. 1 Own Exp.	.03300	.03262	.03263	.00070	.03000	.03492
Occ. 1 Own Exp. Sq.	-.00050	-.00049	-.00049	.00002	-.00057	-.00042
Occ. 1 Other Exp.	.00000	.00019	.00016	.00084	-.00259	.00339
Occ. 1 Other Exp Sq.	.00000	-.00005	-.00005	.00003	-.00016	.00005
Occ. 1 Error Var.	.04000	.04072	.04081	.00082	.03785	.04268
Occ. 2 Int.	9.00000	8.99625	8.99657	.00689	8.97321	9.01843
Occ. 2 Own Exp.	.06700	.06978	.06974	.00112	.06571	.07325
Occ. 2 Own Exp. Sq.	-.00100	-.00107	-.00107	.00003	-.00117	-.00096
Occ. 2 Other Exp.	.02200	.01803	.01802	.00138	.01365	.02383
Occ. 2 Other Exp Sq.	-.00050	-.00041	-.00041	.00005	-.00060	-.00026
Occ. 2 Error Var.	.06250	.06346	.06347	.00089	.06012	.06648
Error Covariance	.00000	-.00256	-.00254	.00104	-.00521	.00003
Error Correlation	.00000	-.05026	-.04989	.02026	-.10128	.00066
SD Eta	400.00	412.45	383.08	90.60	280.71	571.95
PI 1	-2861.22	-2275.14	-2268.75	163.01	-2717.07	-1889.73
PI 2	-717.03	-2515.66	-2382.84	721.67	-4595.73	-1091.64
PI 3	2285.40	4402.92	4295.74	589.58	3152.15	5861.02
PI 4	-790.83	-1783.01	-1745.13	178.52	-2266.54	-1395.11
PI 5	108.03	270.88	268.67	21.85	222.86	326.34
PI 6	-5.24	-13.73	-13.78	1.02	-16.05	-11.15
PI 7	137.86	455.12	357.22	565.73	-988.30	1915.07
PI 8	-5757.05	-6780.05	-6800.54	628.35	-8489.48	-5436.43
PI 9	2835.75	4184.87	4205.19	253.29	3567.70	4795.58
PI10	-437.15	-747.73	-757.16	51.73	-870.11	-624.87
PI11	21.28	40.73	40.61	3.85	32.43	49.02
PI12	3763.28	4108.51	4135.01	304.92	3349.37	4887.98
PI13	-2986.69	-4432.14	-4446.77	226.42	-5031.70	-3965.43
PI14	479.48	897.15	889.32	88.34	762.91	1125.58
PI15	-20.18	-50.66	-49.66	7.71	-69.38	-37.82
PI16	897.49	1455.56	1439.65	84.79	1235.66	1729.37
PI17	-88.54	-316.72	-306.49	63.99	-471.94	-196.24
PI18	-4.20	17.16	16.51	6.42	3.54	31.86
PI19	-57.82	-22.86	-24.30	18.34	-66.50	14.75
PI20	8.92	2.61	2.68	2.08	-2.05	8.04
PI21	-.74	-.01	-.01	.18	-.73	.57

Data Set 2-EMAX

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42768	9.42754	.00486	9.41162	9.44391
Occ. 1 Own Exp.	.03300	.03196	.03195	.00078	.02960	.03488
Occ. 1 Own Exp. Sq.	-.00050	-.00047	-.00047	.00002	-.00054	-.00040
Occ. 1 Other Exp.	.00000	.00099	.00098	.00100	-.00329	.00473
Occ. 1 Other Exp Sq.	.00000	-.00005	-.00006	.00004	-.00019	.00010
Occ. 1 Error Var.	.04000	.03958	.03957	.00063	.03770	.04161
Occ. 2 Int.	9.00000	9.00681	9.00698	.00659	8.98613	9.02580
Occ. 2 Own Exp.	.06700	.06658	.06655	.00096	.06370	.06967
Occ. 2 Own Exp. Sq.	-.00100	-.00100	-.00100	.00003	-.00110	-.00091
Occ. 2 Other Exp.	.02200	.02183	.02179	.00150	.01655	.02691
Occ. 2 Other Exp Sq.	-.00050	-.00055	-.00055	.00006	-.00074	-.00036
Occ. 2 Error Var.	.06250	.06132	.06133	.00083	.05895	.06429
Error Covariance	.00000	-.00139	-.00131	.00126	-.00468	.00227
Error Correlation	.00000	-.02807	-.02657	.02551	-.09369	.04519
SD Eta	400.00	923.32	940.83	86.61	718.30	1078.77
PI 1	-2861.22	-2466.07	-2462.99	196.85	-3016.62	-1911.88
PI 2	-717.03	-1209.73	-1236.76	694.31	-3164.75	851.48
PI 3	2285.40	3829.17	3814.17	613.10	2059.06	5472.48
PI 4	-790.83	-1792.41	-1781.49	214.09	-2399.12	-1163.98
PI 5	108.03	287.78	287.38	31.49	186.53	376.17
PI 6	-5.24	-14.91	-14.92	1.64	-19.50	-9.52
PI 7	137.86	-38.20	-30.46	461.49	-1456.46	1535.78
PI 8	-5757.05	-7533.96	-7458.22	858.56	-10031.17	-4597.69
PI 9	2835.75	4826.89	4791.51	391.18	3691.06	6014.80
PI10	-437.15	-842.73	-838.31	72.96	-1054.19	-617.42
PI11	21.28	44.52	44.38	4.61	31.55	56.75
PI12	3763.28	4801.36	4779.32	387.29	3642.51	5795.23
PI13	-2986.69	-4888.25	-4899.15	365.90	-6130.87	-3960.09
PI14	479.48	898.91	897.14	102.49	647.13	1181.67
PI15	-20.18	-46.37	-46.17	8.02	-71.56	-23.77
PI16	897.49	1468.11	1472.57	178.55	1030.87	1950.73
PI17	-88.54	-244.07	-241.51	66.98	-482.06	-21.60
PI18	-4.20	9.44	9.30	7.03	-14.50	33.02
PI19	-57.82	-45.38	-43.78	24.53	-119.13	28.36
PI20	8.92	4.59	4.36	3.33	-3.60	14.88
PI21	-.74	.17	.20	.52	-1.40	1.90

Data Set 3-EMAX

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42285	9.42295	.00454	9.40759	9.43572
Occ. 1 Own Exp.	.03300	.03311	.03309	.00073	.03083	.03557
Occ. 1 Own Exp. Sq.	-.00050	-.00050	-.00049	.00002	-.00057	-.00043
Occ. 1 Other Exp.	.00000	.00173	.00171	.00081	-.00091	.00429
Occ. 1 Other Exp Sq.	.00000	-.00008	-.00008	.00003	-.00019	.00001
Occ. 1 Error Var.	.04000	.03977	.03982	.00063	.03751	.04146
Occ. 2 Int.	9.00000	9.00106	9.00137	.00636	8.97715	9.02368
Occ. 2 Own Exp.	.06700	.06839	.06832	.00111	.06496	.07171
Occ. 2 Own Exp. Sq.	-.00100	-.00105	-.00105	.00003	-.00115	-.00095
Occ. 2 Other Exp.	.02200	.02162	.02157	.00155	.01741	.02702
Occ. 2 Other Exp Sq.	-.00050	-.00052	-.00051	.00006	-.00070	-.00039
Occ. 2 Error Var.	.06250	.06264	.06260	.00088	.05967	.06558
Error Covariance	.00000	.00114	.00112	.00074	-.00120	.00331
Error Correlation	.00000	.02288	.02258	.01479	-.02423	.06645
SD Eta	400.00	376.46	389.90	47.28	278.09	479.99
PI 1	-2861.22	-1944.13	-1944.49	240.25	-2750.59	-1393.96
PI 2	-717.03	-3566.42	-3641.57	632.92	-4905.05	-1208.37
PI 3	2285.40	5104.83	5176.30	580.72	3548.48	6188.79
PI 4	-790.83	-1972.37	-1972.62	193.83	-2346.73	-1394.77
PI 5	108.03	300.46	301.85	27.10	216.53	354.69
PI 6	-5.24	-15.50	-15.59	1.38	-18.27	-11.36
PI 7	137.86	401.09	398.50	444.25	-737.16	1542.29
PI 8	-5757.05	-6515.94	-6496.85	821.56	-8266.17	-4635.38
PI 9	2835.75	4065.07	4075.50	306.68	3240.18	4776.35
PI10	-437.15	-770.07	-763.72	57.24	-944.15	-612.70
PI11	21.28	44.39	43.77	3.84	36.01	57.24
PI12	3763.28	4114.05	4124.81	272.16	3468.44	4930.80
PI13	-2986.69	-4512.99	-4499.29	275.15	-5444.67	-3631.24
PI14	479.48	1019.16	1006.15	87.81	878.82	1317.33
PI15	-20.18	-63.35	-62.27	7.12	-85.08	-50.90
PI16	897.49	1523.18	1495.38	135.72	1232.71	1952.29
PI17	-88.54	-441.26	-434.25	62.32	-603.78	-325.76
PI18	-4.20	30.27	29.69	6.03	18.01	43.50
PI19	-57.82	16.73	13.96	17.44	-22.74	63.08
PI20	8.92	-1.80	-1.80	2.12	-6.91	3.41
PI21	-.74	-.03	-.03	.22	-.72	.85

Data Set 4-EMAX

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42747	9.42744	.00491	9.41011	9.44805
Occ. 1 Own Exp.	.03300	.03221	.03222	.00075	.02980	.03466
Occ. 1 Own Exp. Sq.	-.00050	-.00048	-.00048	.00002	-.00055	-.00042
Occ. 1 Other Exp.	.00000	.00095	.00097	.00105	-.00244	.00456
Occ. 1 Other Exp Sq.	.00000	-.00007	-.00007	.00004	-.00019	.00008
Occ. 1 Error Var.	.04000	.04023	.04023	.00084	.03763	.04286
Occ. 2 Int.	9.00000	8.99930	8.99922	.00675	8.97572	9.02151
Occ. 2 Own Exp.	.06700	.06884	.06881	.00101	.06540	.07234
Occ. 2 Own Exp. Sq.	-.00100	-.00105	-.00105	.00003	-.00116	-.00096
Occ. 2 Other Exp.	.02200	.02074	.02073	.00162	.01572	.02583
Occ. 2 Other Exp Sq.	-.00050	-.00050	-.00051	.00007	-.00073	-.00030
Occ. 2 Error Var.	.06250	.06099	.06098	.00088	.05796	.06391
Error Covariance	.00000	-.00280	-.00296	.00130	-.00676	.00050
Error Correlation	.00000	-.05648	-.05975	.02603	-.13517	.01020
SD Eta	400.00	1144.95	1151.39	79.45	934.89	1379.81
PI 1	-2861.22	-1706.04	-1717.06	233.52	-2512.23	-810.17
PI 2	-717.03	-3784.52	-3768.08	818.64	-6371.53	-1113.53
PI 3	2285.40	5253.57	5251.48	734.46	2691.14	7727.12
PI 4	-790.83	-2210.30	-2208.41	240.23	-2998.12	-1440.29
PI 5	108.03	348.53	348.93	32.68	256.37	445.07
PI 6	-5.24	-18.08	-18.15	1.60	-22.60	-13.79
PI 7	137.86	21.51	-22.72	536.85	-1480.82	2122.87
PI 8	-5757.05	-6457.47	-6411.95	882.95	-8764.30	-4072.67
PI 9	2835.75	4865.96	4871.22	346.21	3872.25	5736.74
PI10	-437.15	-931.15	-934.17	64.32	-1118.24	-768.99
PI11	21.28	52.34	52.44	4.27	40.51	65.00
PI12	3763.28	4425.14	4415.04	360.98	3281.72	5495.69
PI13	-2986.69	-5819.93	-5785.74	371.54	-6869.68	-4815.46
PI14	479.48	1209.70	1208.65	110.46	940.17	1510.05
PI15	-20.18	-68.59	-69.29	8.70	-93.24	-44.84
PI16	897.49	2055.85	2041.20	183.49	1623.48	2591.71
PI17	-88.54	-413.97	-426.13	77.98	-638.10	-207.09
PI18	-4.20	20.22	20.90	7.69	-1.85	45.56
PI19	-57.82	-53.15	-55.84	27.66	-125.26	20.60
PI20	8.92	6.28	6.57	3.43	-7.08	15.72
PI21	-.74	-.09	-.10	.61	-1.87	2.56

Data Set 5-EMAX

<u>Parameter</u>	<u>True</u>	<u>Mean</u>	<u>Med</u>	<u>SD</u>	<u>Min</u>	<u>Max</u>
Occ. 1 Int.	9.42500	9.42501	9.42512	.00476	9.40964	9.44054
Occ. 1 Own Exp.	.03300	.03326	.03324	.00075	.03050	.03582
Occ. 1 Own Exp. Sq.	-.00050	-.00051	-.00051	.00002	-.00058	-.00043
Occ. 1 Other Exp.	.00000	.00116	.00115	.00091	-.00197	.00402
Occ. 1 Other Exp Sq.	.00000	-.00006	-.00006	.00003	-.00016	.00004
Occ. 1 Error Var.	.04000	.03947	.03949	.00073	.03725	.04221
Occ. 2 Int.	9.00000	9.00383	9.00383	.00614	8.98340	9.02467
Occ. 2 Own Exp.	.06700	.06845	.06845	.00093	.06557	.07166
Occ. 2 Own Exp. Sq.	-.00100	-.00103	-.00103	.00003	-.00113	-.00095
Occ. 2 Other Exp.	.02200	.01861	.01859	.00139	.01488	.02356
Occ. 2 Other Exp Sq.	-.00050	-.00041	-.00041	.00006	-.00059	-.00026
Occ. 2 Error Var.	.06250	.06379	.06378	.00088	.06117	.06646
Error Covariance	.00000	.00007	.00009	.00097	-.00267	.00325
Error Correlation	.00000	.00154	.00180	.01929	-.05266	.06476
SD Eta	400.00	743.69	734.47	54.89	617.31	905.40
PI 1	-2861.22	-2254.28	-2243.22	202.82	-2796.61	-1696.53
PI 2	-717.03	-2534.70	-2553.71	751.72	-5107.52	-432.79
PI 3	2285.40	4384.96	4382.24	669.77	2530.72	6623.84
PI 4	-790.83	-1805.12	-1771.68	225.74	-2579.87	-1209.68
PI 5	108.03	277.95	270.78	32.42	201.31	379.86
PI 6	-5.24	-14.21	-13.82	1.68	-19.33	-10.70
PI 7	137.86	277.55	290.25	434.78	-855.72	1563.31
PI 8	-5757.05	-6844.18	-6745.88	739.95	-9144.72	-4810.17
PI 9	2835.75	4314.75	4264.33	358.08	3551.80	5401.98
PI10	-437.15	-757.24	-749.31	77.59	-976.35	-574.41
PI11	21.28	40.23	39.86	5.30	27.94	54.59
PI12	3763.28	4279.90	4252.99	341.38	3303.59	5357.59
PI13	-2986.69	-4543.45	-4514.26	371.71	-5588.27	-3725.08
PI14	479.48	829.44	821.11	117.22	589.23	1176.68
PI15	-20.18	-41.67	-40.59	9.26	-67.85	-23.37
PI16	897.49	1438.60	1441.07	165.12	1001.62	2073.02
PI17	-88.54	-186.41	-178.87	74.59	-394.69	-23.84
PI18	-4.20	2.29	1.48	7.02	-13.99	22.25
PI19	-57.82	-78.93	-82.33	20.38	-124.39	-11.91
PI20	8.92	9.31	9.55	2.41	1.72	15.30
PI21	-.74	-.14	-.12	.43	-1.58	1.31

Table 6

OLS Wage Equation Parameter Estimates for Data Generated with True Future Component (standard deviations in parentheses)*

Data Set	Occupation One					Occupation Two					Wage Error SDs	
	Intercept	Own Exp.	Own Exp. Squared	Other Exp.	Other Exp. Squared	Intercept	Own Exp.	Own Exp. Squared	Other Exp.	Other Exp. Squared	Occ. 1	Occ. 2
Actual	9.42500	0.03300	-0.00050	0.00000	0.00000	9.00000	0.06700	-0.00100	0.02200	-0.00050	0.20000	0.25000
1-EMAX	9.50638 (0.00470)	0.02627 (0.00074)	-0.00039 (0.00002)	0.02123 (0.00107)	-0.00054 (0.00005)	9.17518 (0.00635)	0.05938 (0.00094)	-0.00086 (0.00003)	0.00796 (0.00169)	0.00031 (0.00009)	0.17981	0.22853
2-EMAX	9.50716 (0.00475)	0.02662 (0.00074)	-0.00038 (0.00002)	0.02098 (0.00104)	-0.00053 (0.00005)	9.17977 (0.00648)	0.05829 (0.00095)	-0.00083 (0.00003)	0.00907 (0.00188)	0.00024 (0.00011)	0.17890	0.22594
3-EMAX	9.49973 (0.00469)	0.02763 (0.00071)	-0.00040 (0.00002)	0.02251 (0.00106)	-0.00059 (0.00005)	9.17000 (0.00658)	0.05893 (0.00095)	-0.00086 (0.00003)	0.01066 (0.00183)	0.00017 (0.00010)	0.17999	0.22882
4-EMAX	9.50622 (0.00474)	0.02653 (0.00072)	-0.00039 (0.00002)	0.02164 (0.00106)	-0.00054 (0.00005)	9.16782 (0.006414)	0.06002 (0.000945)	-0.00088 (0.00003)	0.00962 (0.00177)	0.00023 (0.00010)	0.18052	0.22606
5-EMAX	9.50534 (0.00475)	0.02773 (0.00074)	-0.00042 (0.00002)	0.02081 (0.00100)	-0.00055 (0.00004)	9.17511 (0.00629)	0.05929 (0.00094)	-0.00085 (0.00003)	0.00808 (0.00169)	0.00025 (0.00009)	0.17916	0.22986

*Results from simple OLS regression using only observed wages and experience.

Table 7

Analysis of Wealth Loss when Estimated
Polynomial Approximation is used in Place of True Future Component*

Data Set	Using True EMAX**	Using Approximate EMAX**						
	Mean Wealth***	Mean Wealth***	Mean Dollar Loss	Mean Percent Loss	Aggregate Choice Agreement	Percent with 0-35 Agreements	Percent with 36-39 Agreements	Percent Choosing Same Path
1-EMAX	354534.72	354345.25	189.47	0.05%	94.10%	9.60%	39.40%	51.00%
2-EMAX	354912.59	354529.61	382.98	0.11%	94.60%	7.20%	38.40%	54.40%
3-EMAX	356430.12	356175.54	254.58	0.07%	93.90%	10.20%	39.20%	50.60%
4-EMAX	355723.35	355589.21	134.14	0.03%	95.00%	6.80%	35.80%	57.40%
5-EMAX	355438.58	354984.04	454.54	0.12%	93.90%	8.60%	39.00%	52.40%

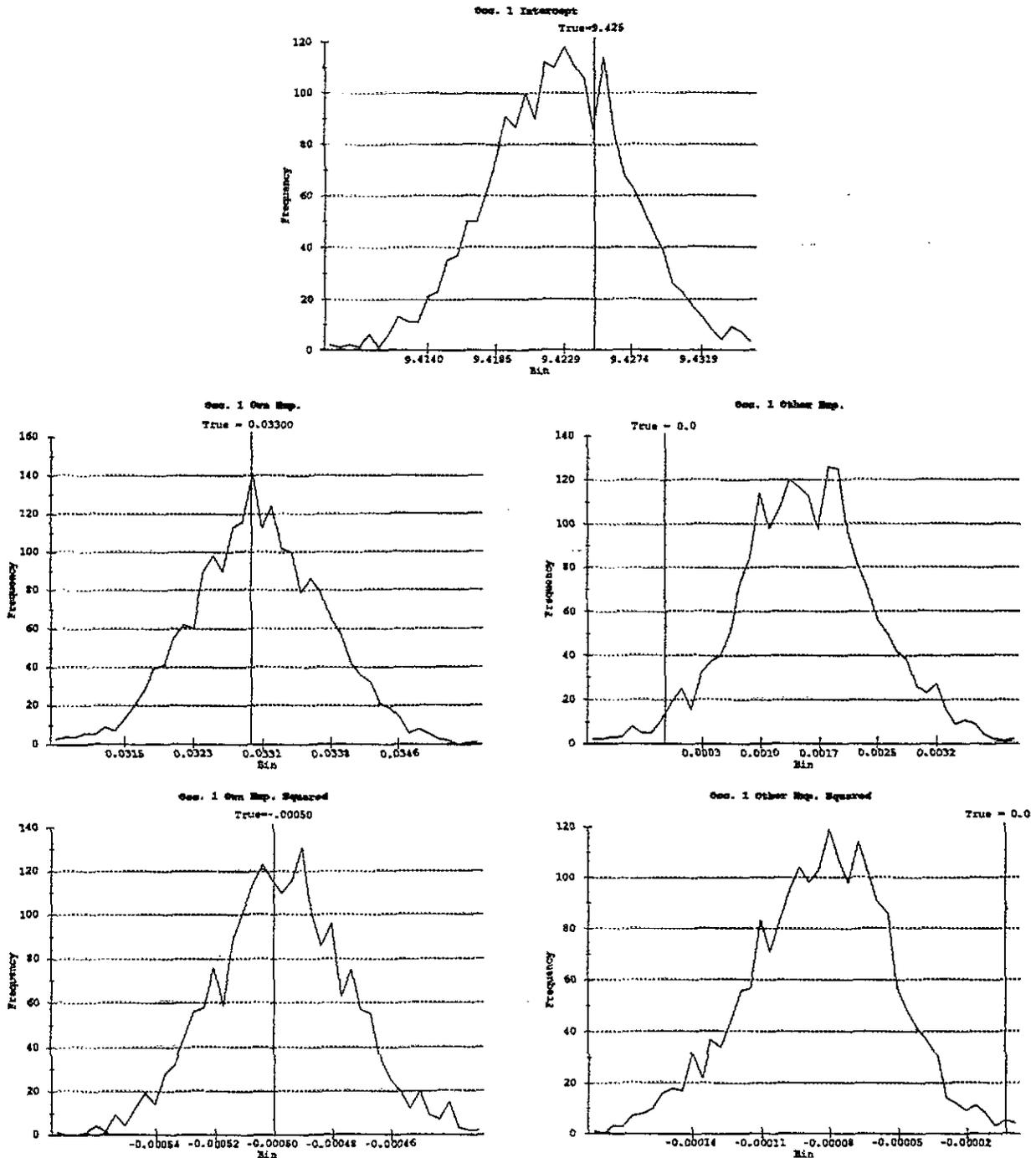
* Polynomial parameter values are set to the mean of their respective empirical posterior distributions.

** Each simulation includes 500 agents that live for exactly 40 periods.

*** "Mean wealth" is the equal-weight sample average of discounted streams of ex-post lifetime earnings.

Figure 1

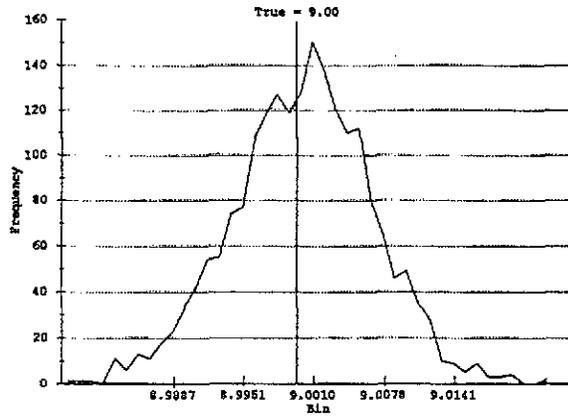
Empirical Posterior Distributions* of Structural Parameters**



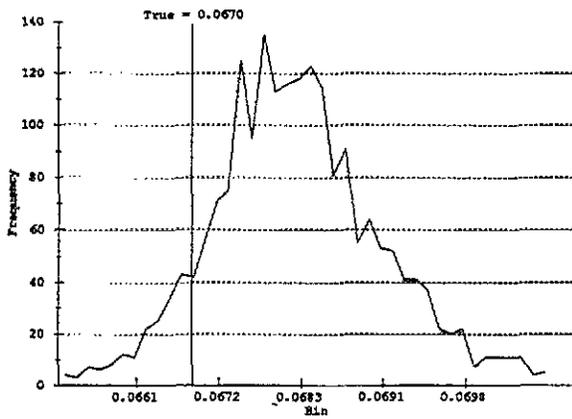
* Based on last 2000 draws of the data set
3-EMAX estimation routine.

** Each graph contains five tick marks; the middle indicates the posterior mean and the others one and two standard deviations from mean.

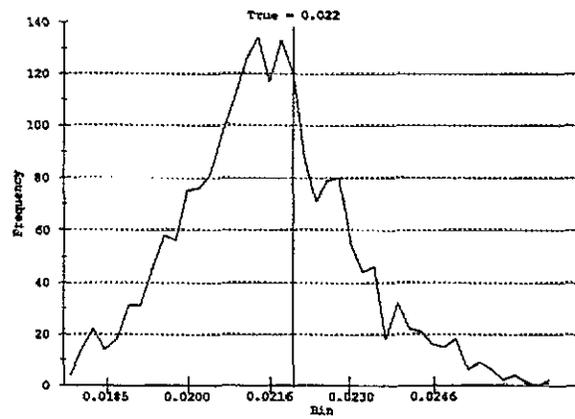
Occ. 2 Intercept



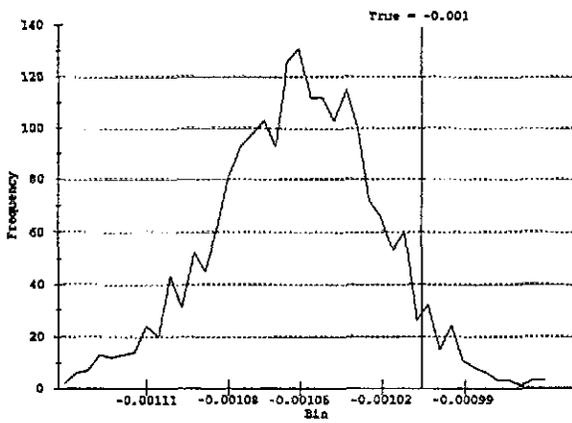
Occ. 2 Own Exp.



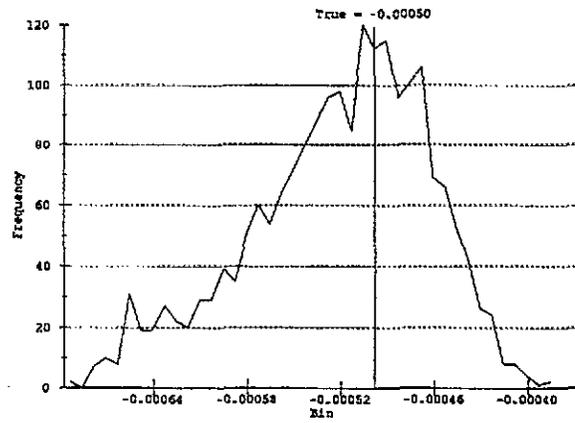
Occ. 2 Other Exp.



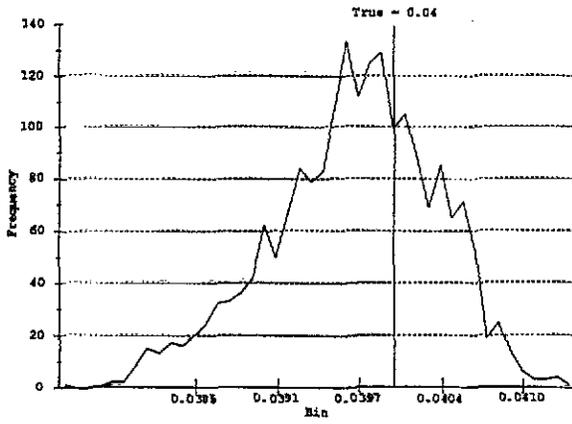
Occ. 2 Own Exp. Squared



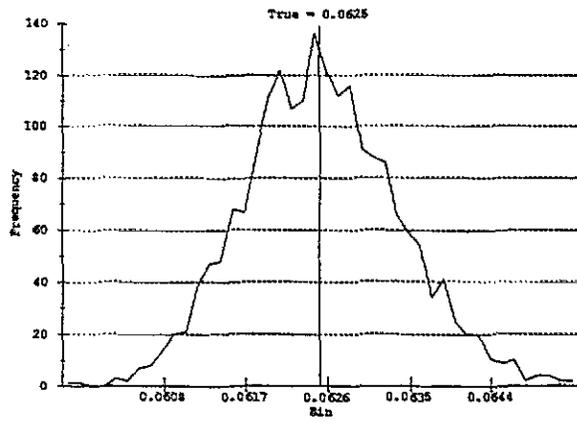
Occ. 2 Other Exp. Squared



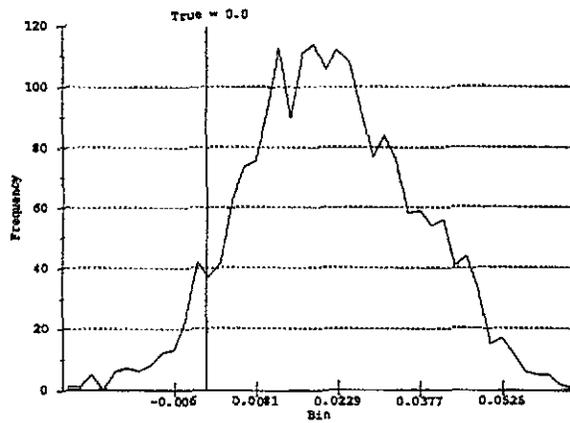
Occ. 1 Wage Error Variance



Occ. 2 Wage Error Variance



Wage Error Correlation



standard Deviation of η_{2t}

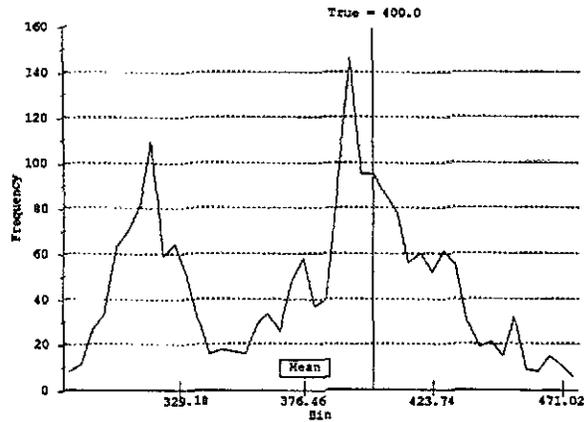
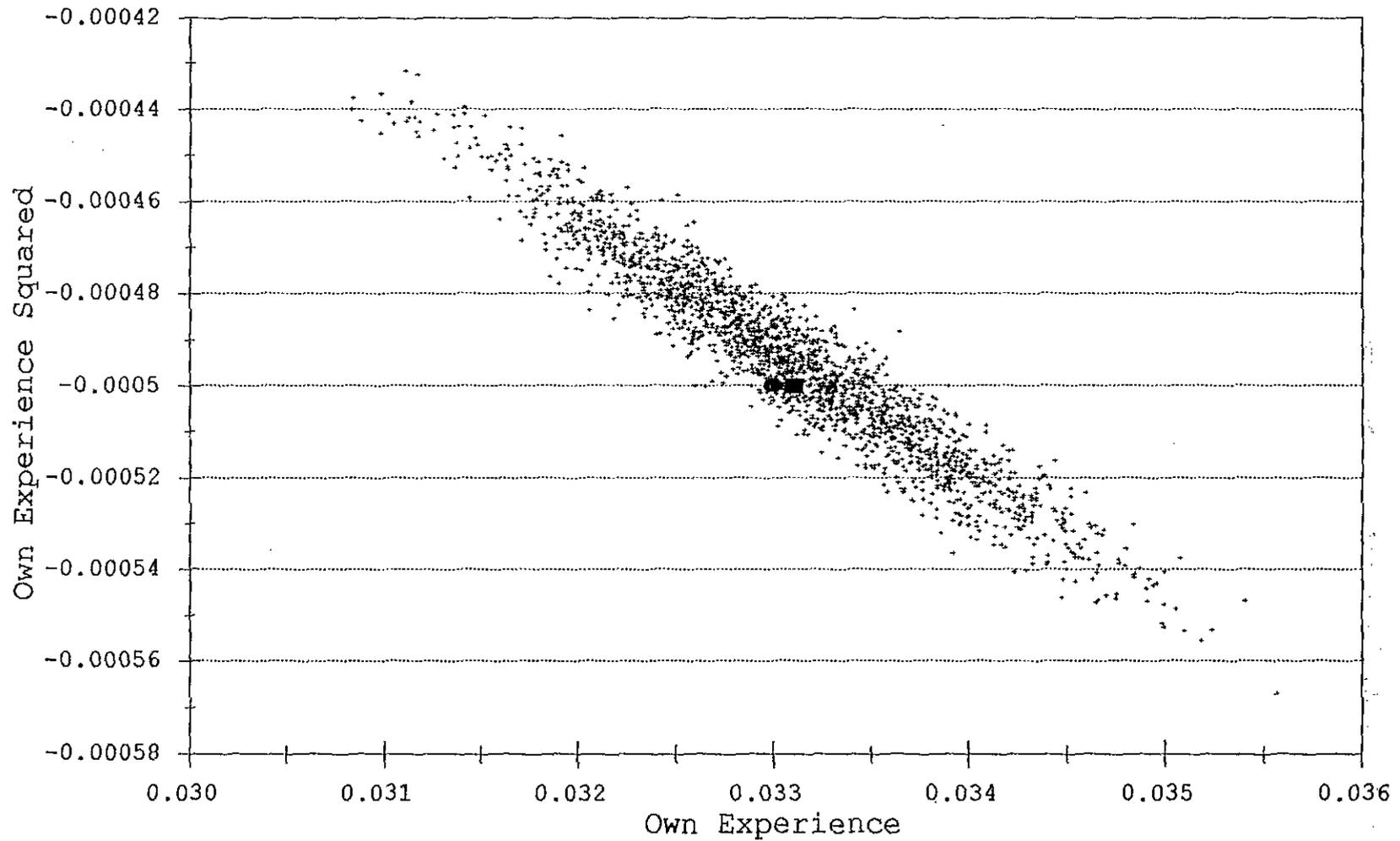


Figure 2A

Joint Density* of Own Experience
and Own Experience Squared for Occ. 1

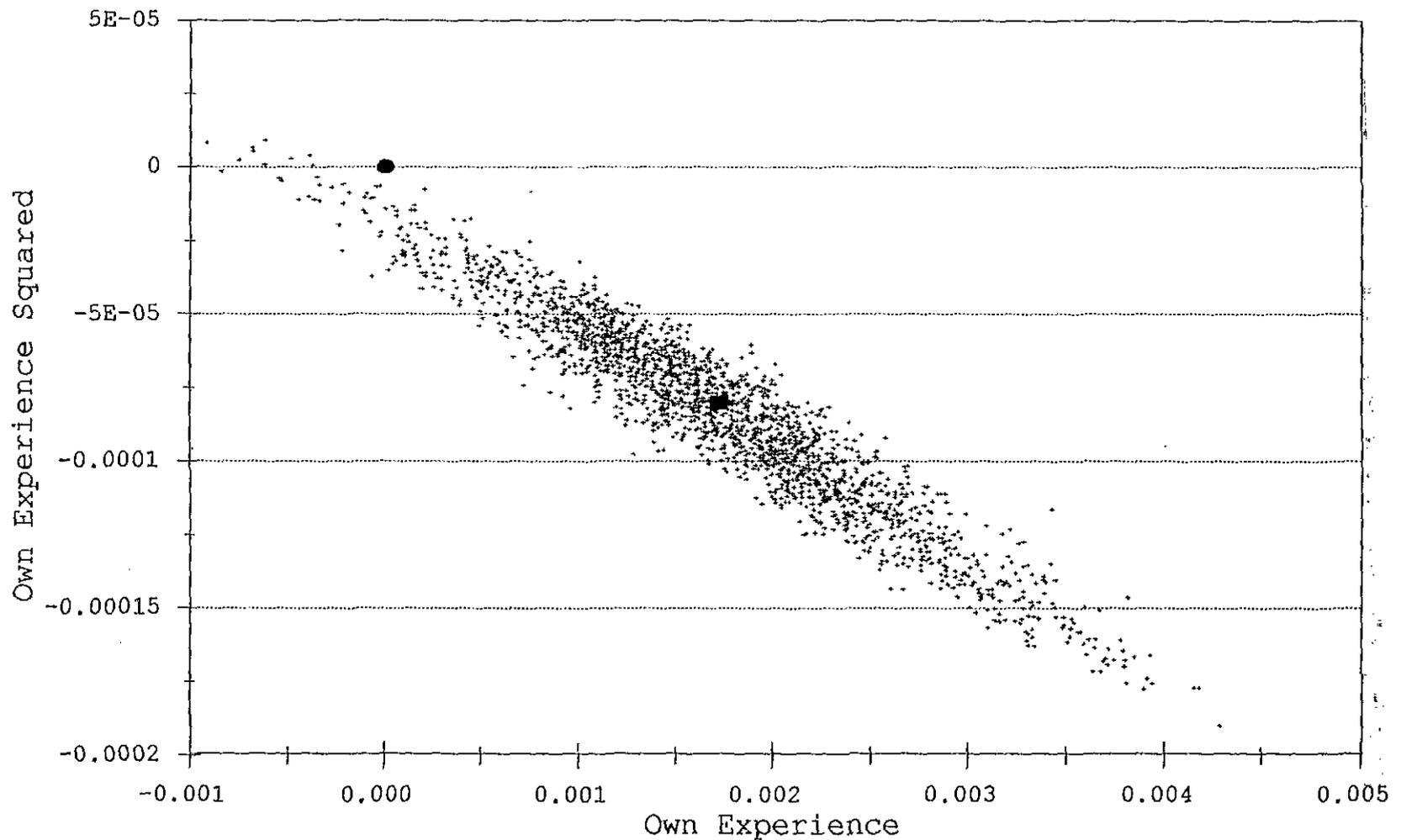


*Results from Data Set
3-EMAX estimation.

● True ■ Mean

Figure 2B

Joint Density* of Other Experience
and Other Exp. Squared for Occ. 1



*Results from Data Set
3-EMAX estimation.

● True ■ Mean

Figure 3A

True Difference in Future Components
at All States in First Five Periods

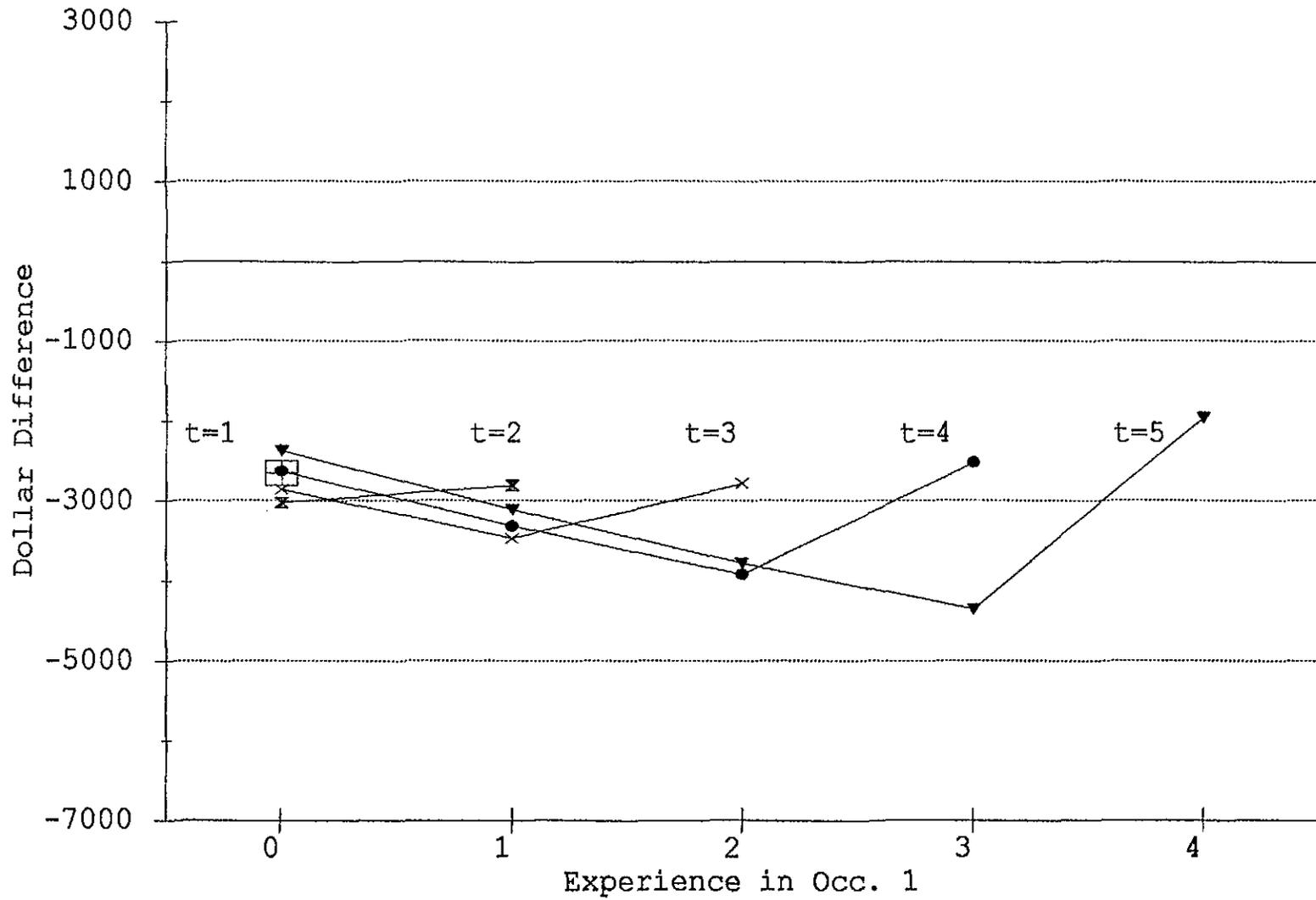


Figure 3B

Estimated Difference*
in Future Components at All
States in First Five Periods

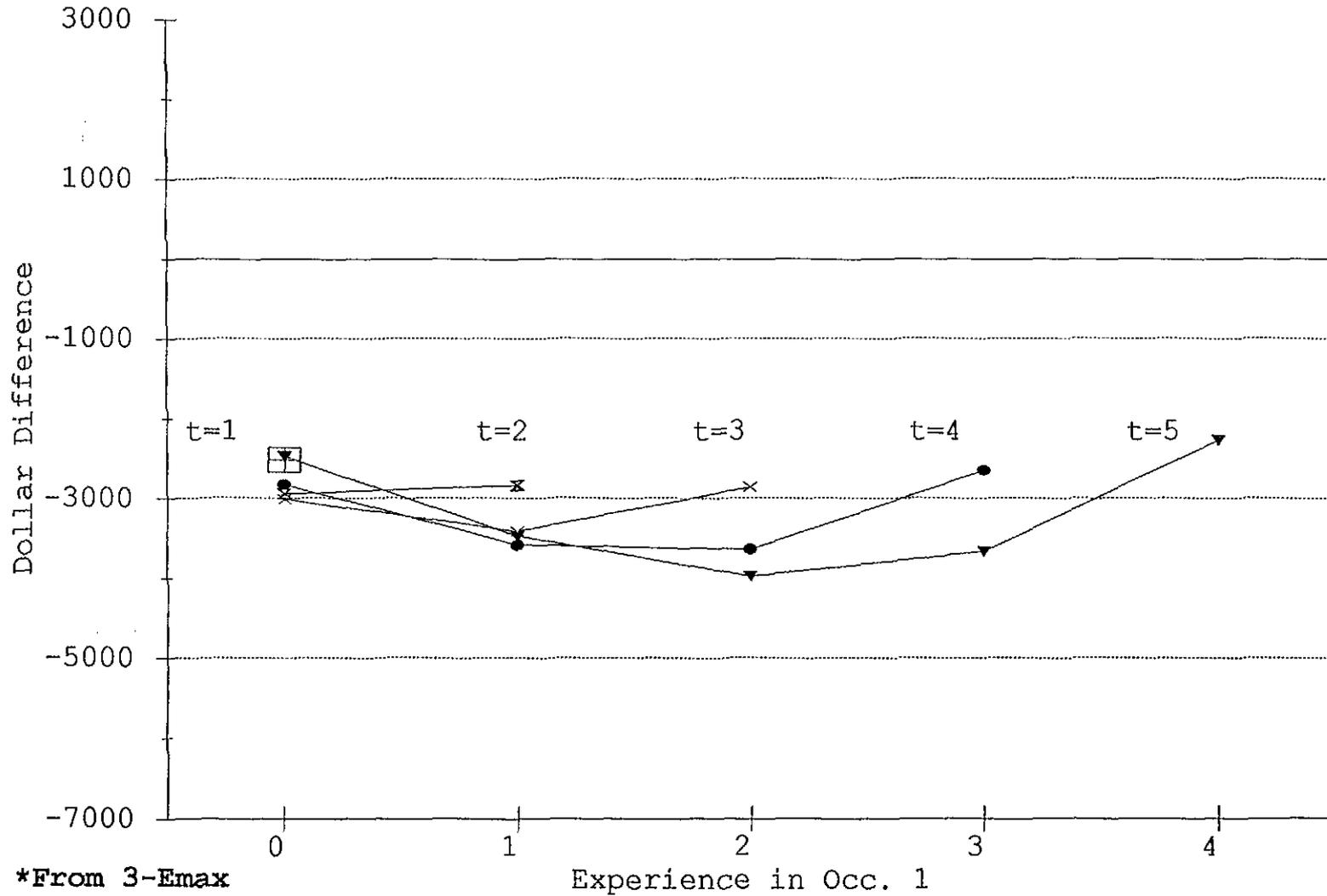


Figure 4A

True Difference in Future Components
at All States in First Ten Periods

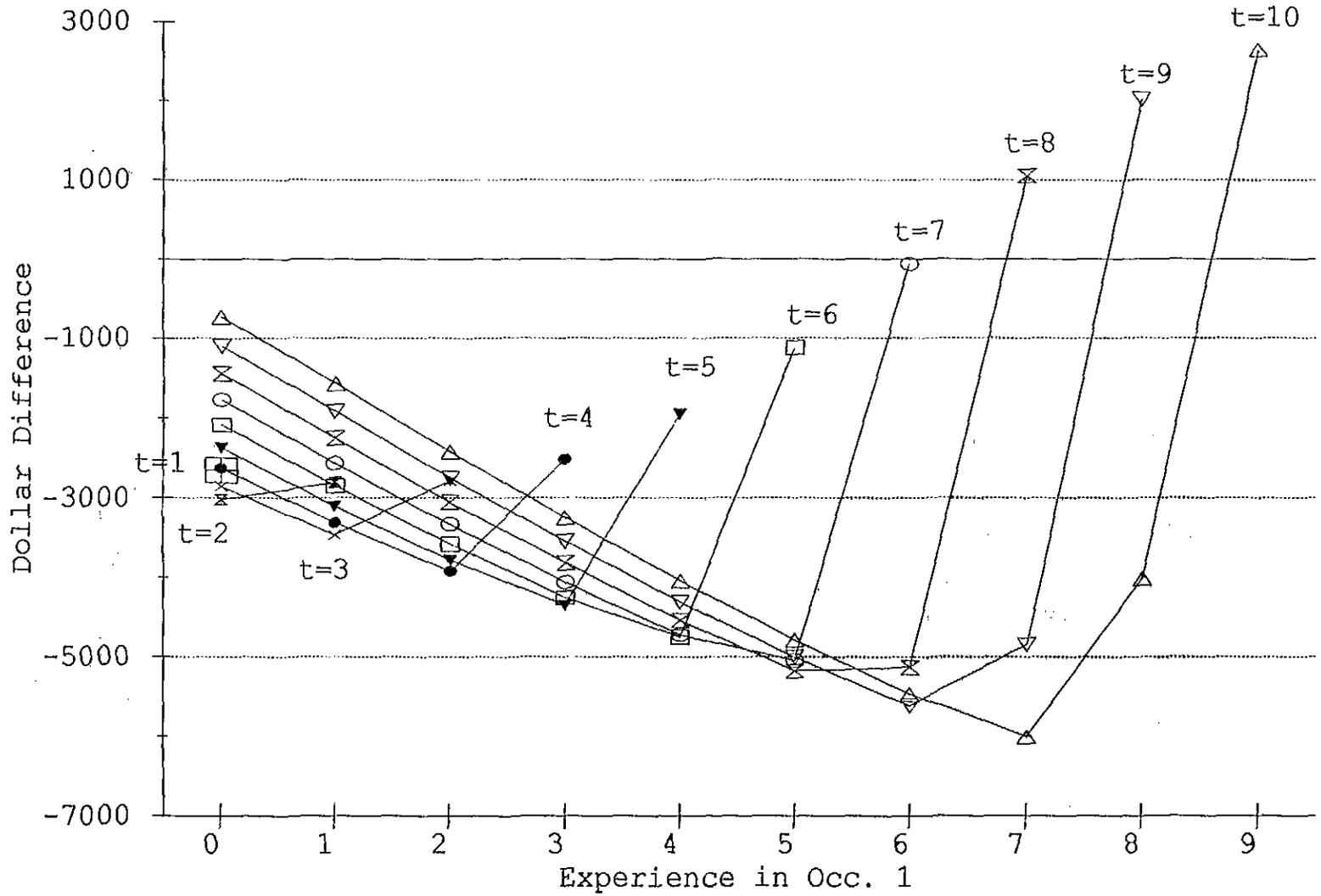
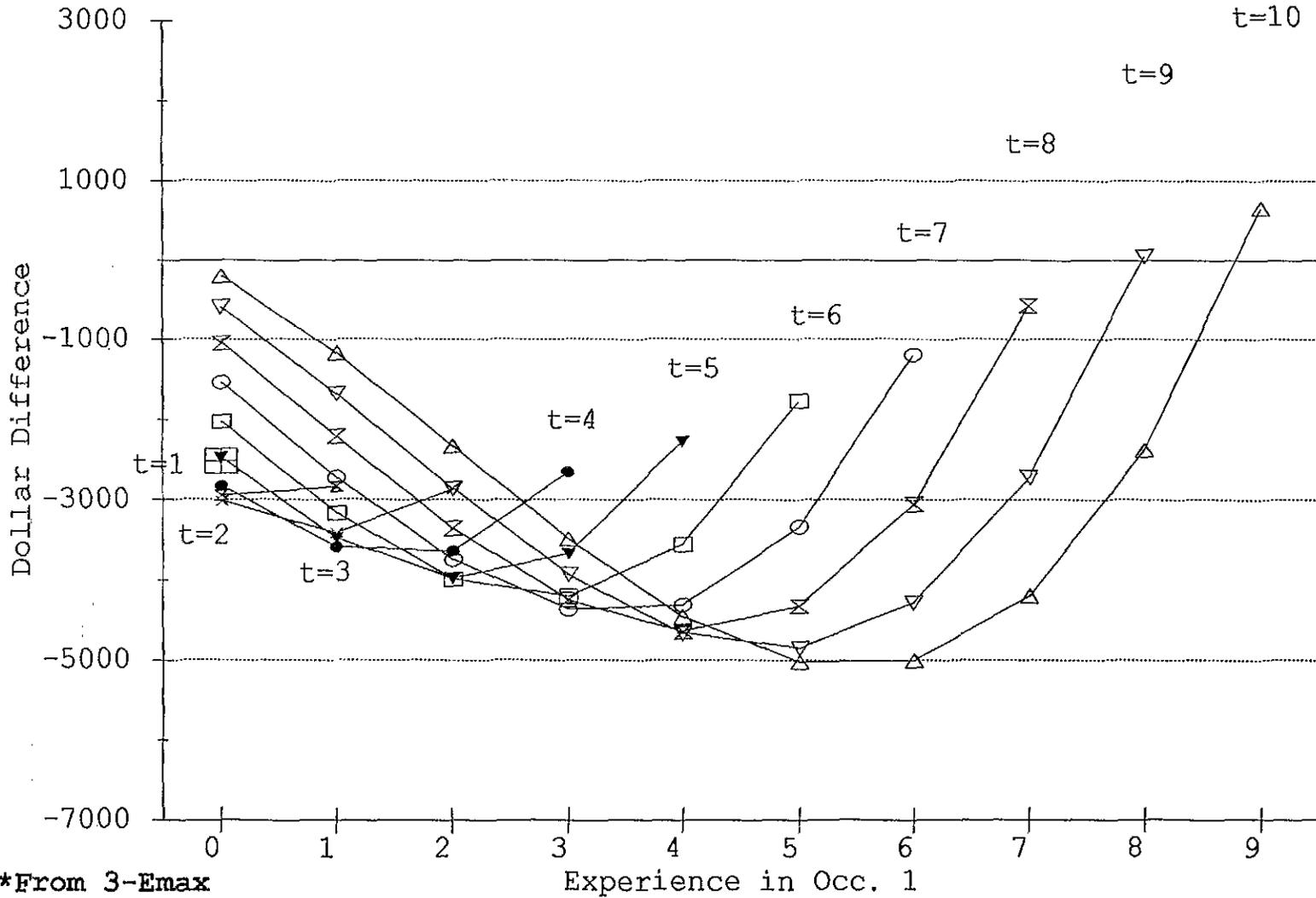


Figure 4B

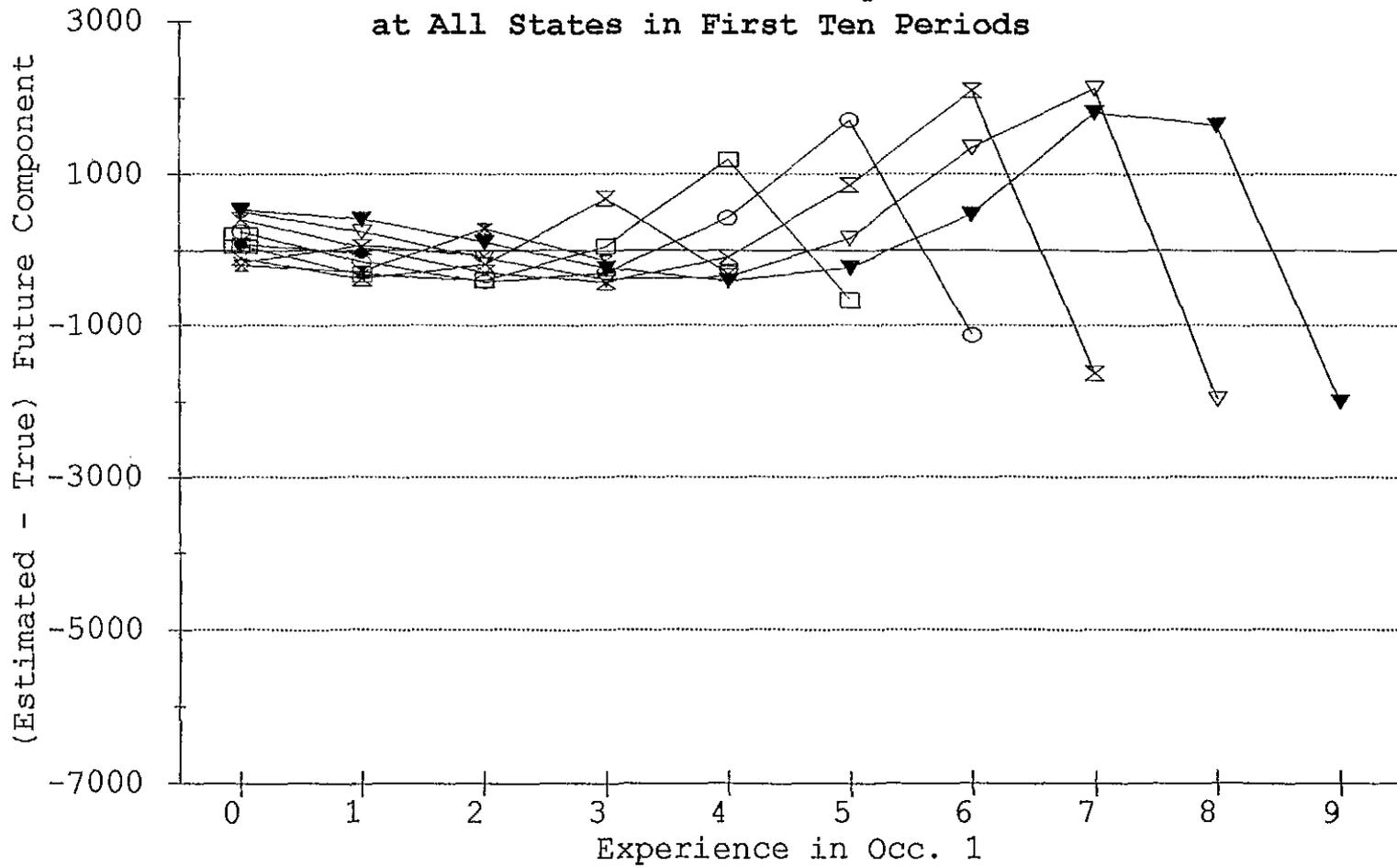
Estimated Difference*
in Future Components at All
States in First Ten Periods



*From 3-Emax
Estimation.

Figure 4C

Dollar Error in Estimated
Difference* of Future Components
at All States in First Ten Periods



*From 3-Emax
Estimation.

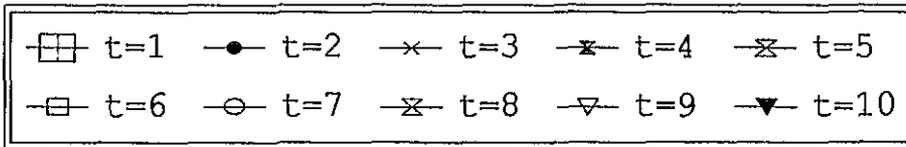
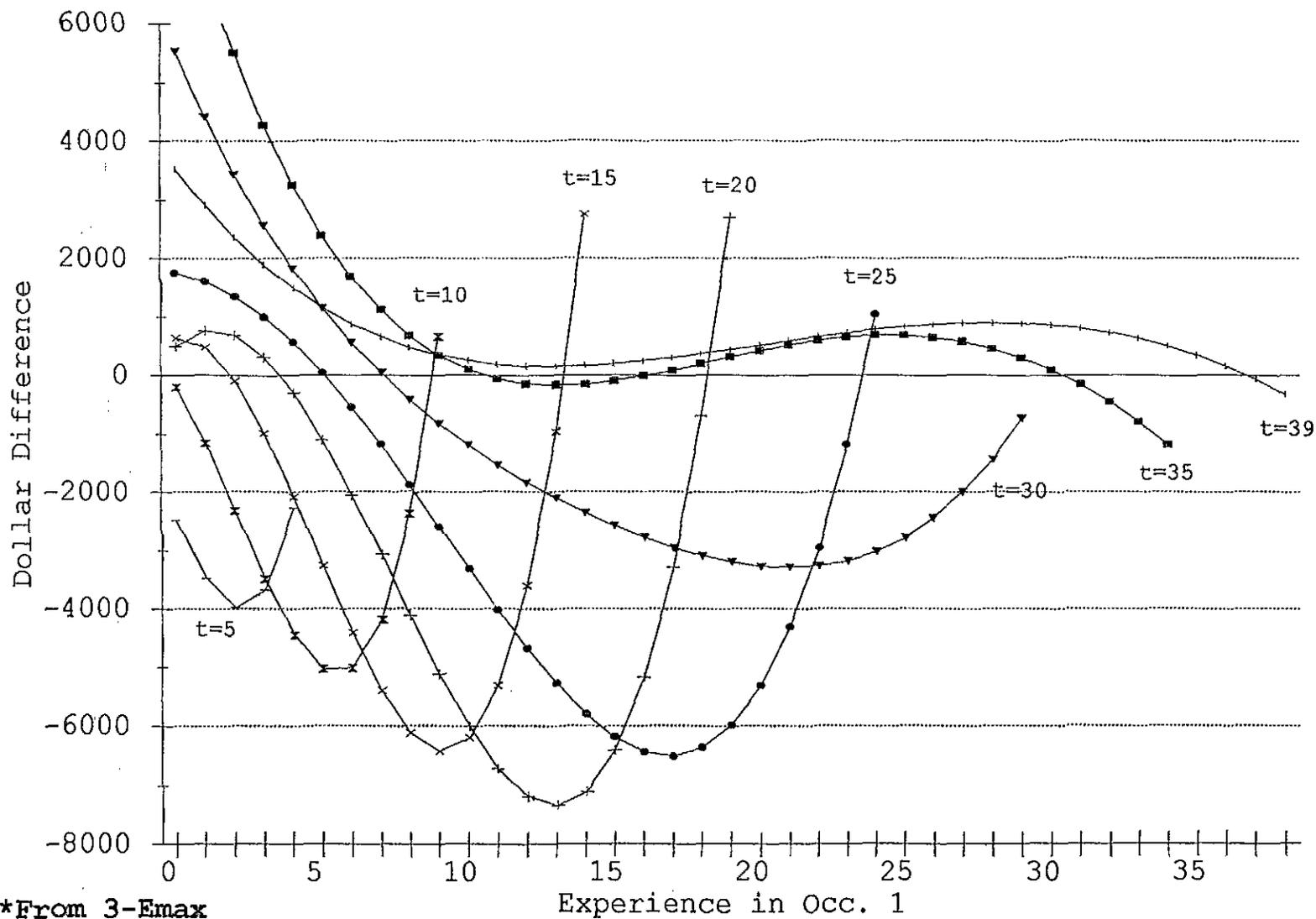


Figure 5B
 Estimated Difference* in Future
 Components at Various State Points

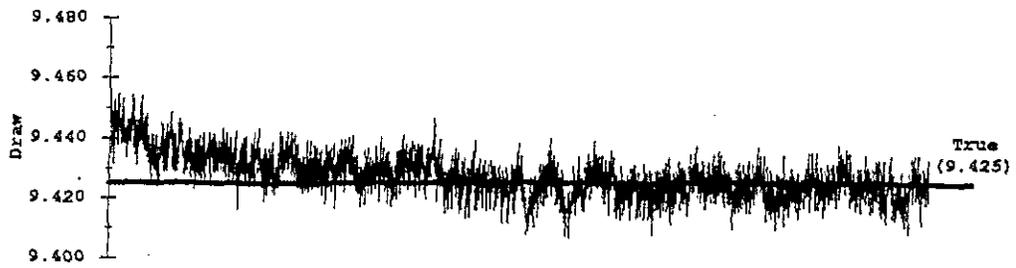


*From 3-Emax
 Estimation.

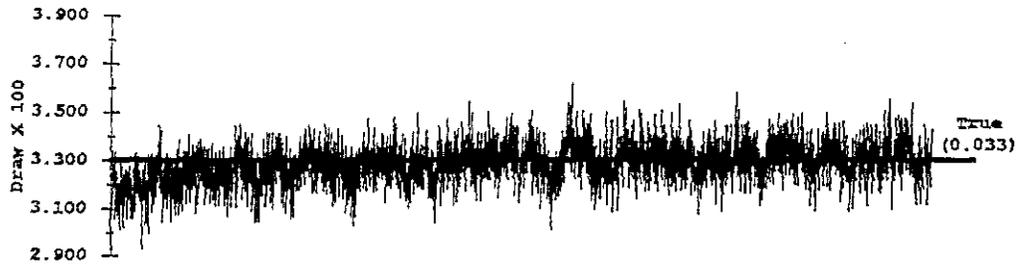
Figure 6

Gibbs Sampler Convergence*

Occ. 1 Intercept



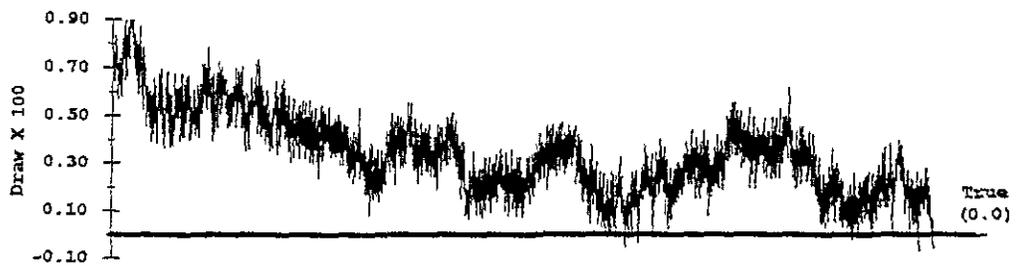
Occ. 1 Own Exp.



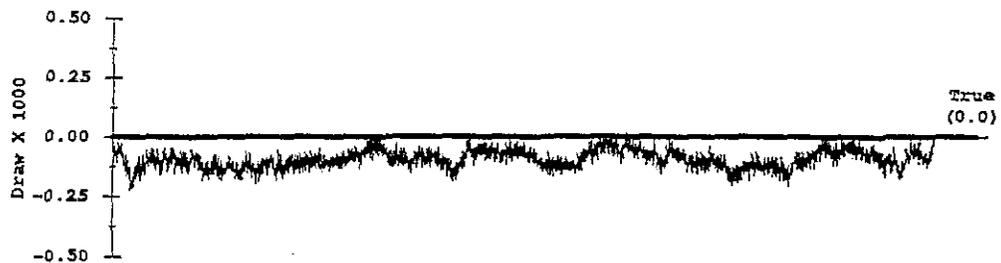
Occ. 1 Own Exp. Sq.



Occ. 1 Other Exp.

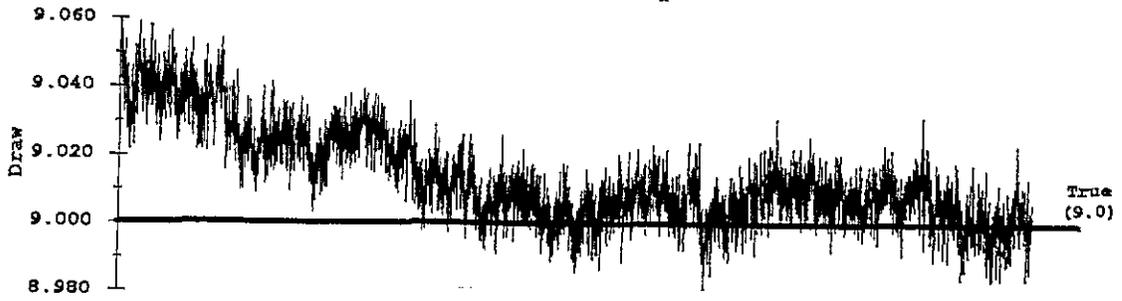


Occ. 1 Other Exp. Sq.

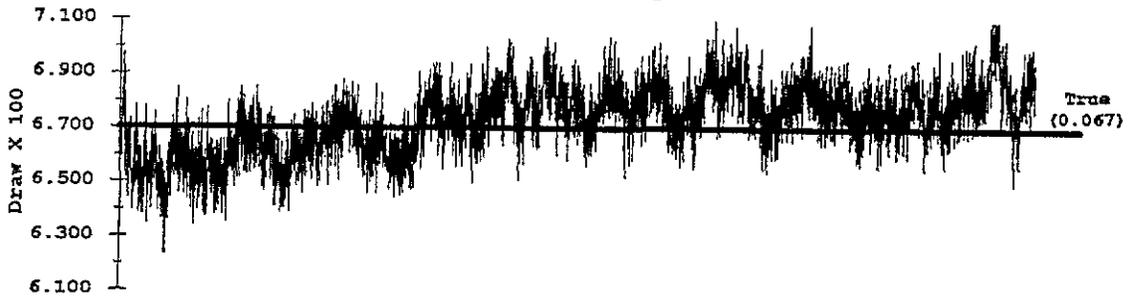


* Results from data set 3-Emax estimation routine.

Occ. 2 Intercept



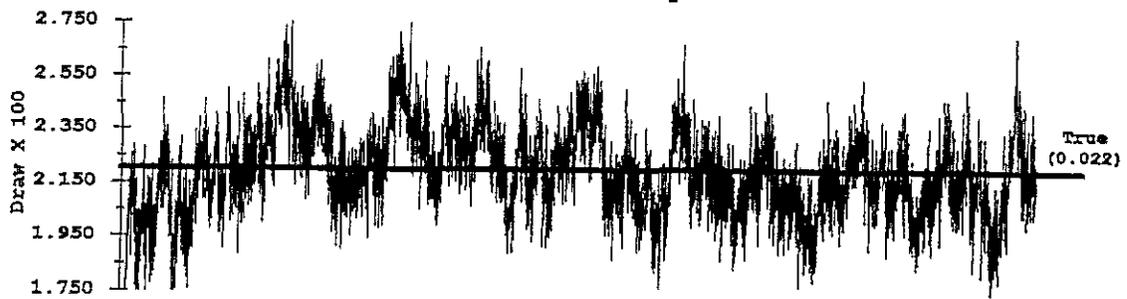
Occ. 2 Own Exp.



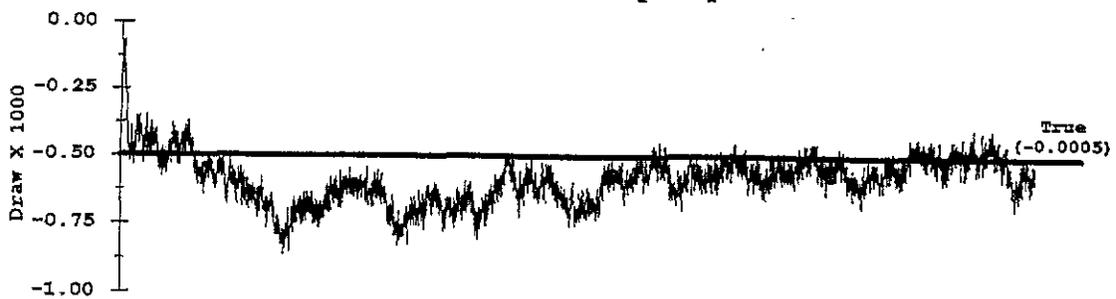
Occ. 2 Own Exp. Sq.



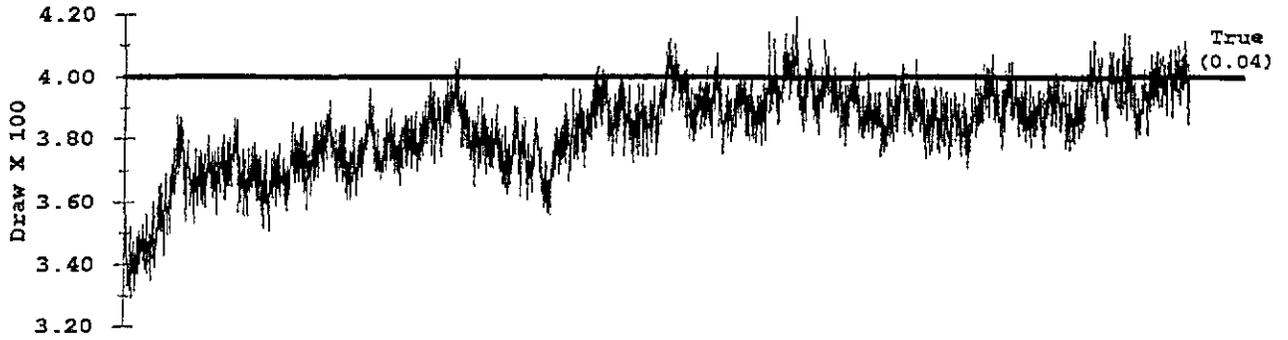
Occ. 2 Other Exp.



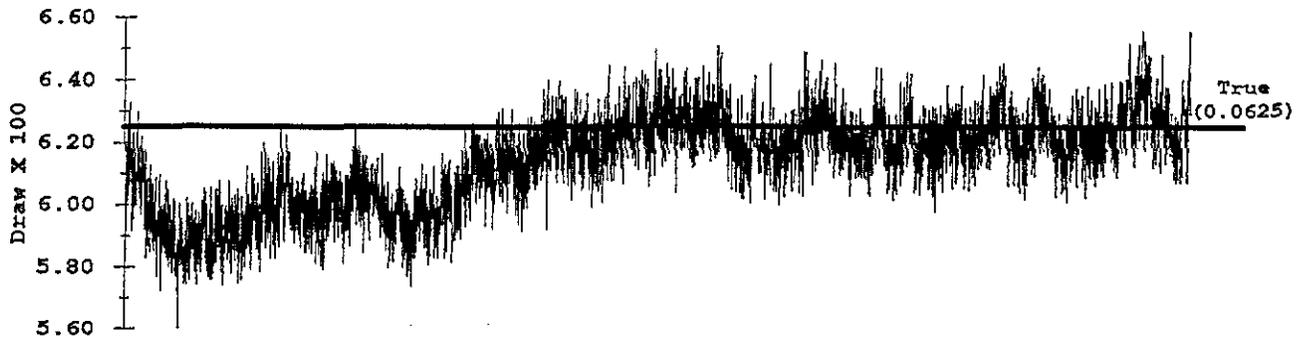
Occ. 2 Other Exp. Sq.



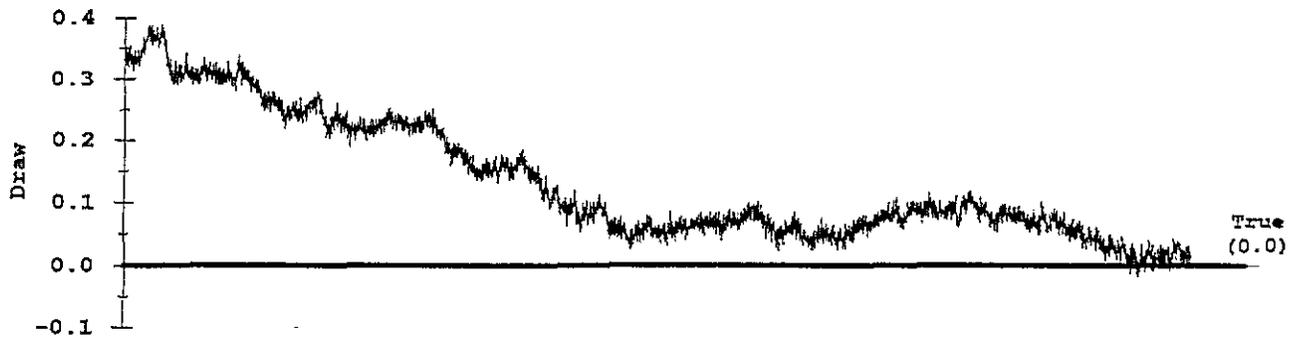
Occ. 1 Wage Error Variance



Occ. 2 Wage Error Variance



Wage Error Variance Correlation



Standard Deviation of Eta

