

Economics Department
Working Papers in Economics

Boston College

Year 1999

Estimating the Rational Expectations
Model of Speculative Storage: A Monte
Carlo Comparison of Three Simulation
Estimators

Alexander Michaelides
Princeton,

Serena Ng
Boston College,

ESTIMATING THE RATIONAL EXPECTATIONS MODEL OF SPECULATIVE STORAGE: A MONTE CARLO COMPARISON OF THREE SIMULATION ESTIMATORS

Alexander Michaelides*
University of Cyprus

Serena Ng†
Boston College

April 1999

Abstract

The non-negativity constraint on inventories imposed on the rational expectations theory of speculative storage implies that the conditional mean and variance of commodity prices are nonlinear in lagged prices and have a kink at a threshold point. In this paper, the structural parameters of this model are estimated using three simulation based estimators. In a Monte-Carlo experiment, the finite sample properties of the Simulated Methods of Moments estimator of Duffie and Singleton (1993), the Indirect Inference estimator of Gourieroux, Monfort and Renault (1993), and the Efficient Method of Moments estimator of Gallant and Tauchen (1996) are assessed. Exploiting the invariant distribution implied by the theory allows us to evaluate the error induced by simulations. Our results show that the estimators differ in their sensitivity to the sample size, the number of simulations, choice of auxiliary models, and computation demands. For some estimators, the test for overidentifying restrictions exhibit significant size distortions in small samples. Overall, while the simulation estimators have small bias, they are less efficient than pseudo maximum likelihood (PMLE). Hence for the small sample sizes considered, the simulation estimators are still inferior to the PMLE estimates in a mean-squared sense.

Keywords: simulation estimators, indirect inference, simulated method of moments, which moments to match, commodity prices.

JEL Classification: C15, B4, G1, Q1.

*The first author thanks the Lynde and Harry Bradley Foundation for financial support. E-mail: alexm@ucy.ac.cy.

†The second author acknowledges grants from the Social Science and Humanities Research Council of Canada (SSHRC) and the Fonds de la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR). This research project was also partially funded under NSF grant SBR 9507809. We would like to thank Angus Deaton for initiating the project and for his comments on earlier drafts. Helpful suggestions from Bo Honoré, Gregory Chow and Mark Watson are also gratefully acknowledged. Remaining errors are our own.

Correspondence: Department of Economics, Boston College, Chestnut Hill, MA 02467. Email: serena.ng@bc.edu

1 Introduction

In recent years, there has been a proliferation of economic models that compute policy functions as solutions to dynamic programming problems. In partial equilibrium models of consumption, for instance, researchers have derived policy functions relating consumption to a state variable denoting liquid wealth. These policy functions are then used to analyze consumption behavior under liquidity constraints and the role of buffer stock savings¹. On the investment side, a number of dynamic programs have been used to analyze investment behavior under uncertainty, while others have examined the investment behavior of liquidity constrained firms². In general equilibrium models of business cycles, dynamic programming has become the major toolkit for analyzing the evolution of the endogenous variables. The complexity of these models, however, has led to only a few attempts at structural parameter estimation, and when such estimation is attempted, limited information techniques such as Generalized Methods of Moments (GMM) are applied to Euler equations.³ Recent developments of computation intensive estimators have opened up a new horizon for the estimation of non-linear dynamic models with rational expectations. Ohanian, Violante, Krusell, and Ríos-Rull (1997), for example, have applied simulation based estimators to non-linear production functions with latent factors, while Gourinchas and Parker (1996) employed the Simulated Method of Moments (SMM) to estimate a life cycle model with buffer stock saving. In the industrial organization literature, models utilizing dynamic programming techniques have been estimated by Pakes (1994) and Rust (1994). Important contributions have also been made by Rust (1987) and Hotz and Miller (1993), among many others.

While simulation based estimators are solidly grounded in theory, little is known about their properties in practice. The purpose of this paper is to investigate the finite sample properties of three simulation based estimators in the context of the speculative storage model under rational expectations. The commodity price model provides an interesting economic environment for the exercise for four reasons. First, the model implies that commodity prices have time varying conditional mean and variance, and that the price function is kinked at a threshold point. These price dynamics are sufficiently complex to put the three estimators to a challenge. Second, since the properties of the Pseudo-Maximum Likelihood (PMLE) estimator for estimating the speculative storage model have been analyzed in Deaton-Laroque (1996), the properties of the three estimators can be compared using the PMLE, a non-simulation estimator, as the benchmark. Third, as will become clear, the objective function of two of the estimators involves approximating an

¹See Deaton (1991), Carroll (1992), Carroll (1997), Carroll and Samwick (1997, 1998), Ludvigson (1996), Gourinchas and Parker (1996), and Ludvigson and Michaelides (1999).

²See, for instance, Dixit and Pindyck (1994), and Gross (1994).

³Exceptions are Altug (1989), and Mcgratten, Rogerson and Wright (1992).

expectational term by Monte Carlo integration. However, the commodity prices derived from the speculative storage model follow a renewal process and have an invariant distribution. Thus, we can also estimate the parameters by using the invariant distribution to evaluate the moments of interest. This is useful because using the invariant distribution amounts to executing an infinite number of simulations to approximate the expectations. Comparison of the estimates based on Monte Carlo integration with those based on the invariant distribution allows us to isolate the error induced by Monte Carlo integration. The exercise also serves as a guide to the choice of the number of simulated data points necessary when the integration is computed by simulation (which is used in most practical applications). Fourth, the model resembles consumption and investment models in which agents are subject to occasionally binding constraints. Our results therefore shed light on the usefulness of the estimators for a broader range of economic problems.

The paper is structured as follows. The estimators are defined in Section 2, and in Section 3 we estimate a simple MA(1) model to gain some understanding of potential pitfalls before implementing the more complicated non-linear rational expectations model. The commodity price model is described in Section 4. The setup of the Monte Carlo experiments is detailed in Section 5, and the results reported in Section 6. In section 7, the simulation bias of the estimators is assessed. Section 8 concludes.

2 The Simulation Estimators

Consider the dynamic model given by

$$y_t = f(\theta, y_{t-1}, x_t, u_t), t = 1, \dots, T \quad (1)$$

$$u_t = \rho u_{t-1} + e_t, \quad (2)$$

where θ is a p dimensional vector of structural parameters, x_t is a vector of exogenous variables, y_t is the endogenous variable, $\{e_t\}$ is an unobserved white noise process with distribution G whose parameters are unknown. The objective is to estimate θ when $f(\cdot)$ does not have a tractable form, but the data for y_t can be simulated from an assumed structural model for a given set of structural parameters, θ .

We consider i) the Indirect Inference Estimator (IND) of Gourieroux, Monfort, and Renault (1993), ii) the Simulated Method of Moments Estimator (SMM) of Duffie and Singleton (1993), and iii) the Efficient Method of Moments Estimator (EMM) developed in Gallant and Tauchen (1996). The estimators require simulating data from an assumed structural model with errors that have the same distribution as, but are independent of, the errors e_t . In all three cases, $\hat{\theta}$ is

determined (empirically) as:

$$\hat{\theta} = \text{Argmin}_{\theta} D' \Omega D. \quad (3)$$

The three estimators differ in the choice of D and the weighting matrix Ω . Under regularity conditions given in the respective papers, $\sqrt{T}(\hat{\theta} - \theta) \rightarrow N(0, W)$. Our interest is in evaluating the finite sample properties of these estimators based upon the optimal weighting matrix, denoted Ω^* .

We concentrate on the case where there are no exogenous variables and $\rho = 0$. Let Y_t and \tilde{Y}_t denote the observations at time t of the actual and simulated endogenous variables, respectively. Let T be the sample size of the observed series. In the case of EMM and IND, we simulate H paths, each of length T . In the case of SMM, it is more convenient to simulate a total of $T \cdot H$ data points. For the latter, let $Y_{[T]}$ and $\tilde{Y}_{[TH]}$ denote the vectors of actual and simulated endogenous variables of length T and TH , respectively. We have:

SMM

$$\begin{aligned} D &= \left(\frac{1}{T} \sum_{t=1}^T m(Y_t) - \frac{1}{TH} \sum_{t=1}^{TH} m(\tilde{Y}_t) \right), \\ \Omega^* &= I_0^{-1} = \lim_{T \rightarrow \infty} \text{Var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T m(Y_t) \right)^{-1}, \\ W_H^* &= \left(1 + \frac{1}{H} \right) \left(E \left[\frac{\partial m(\tilde{Y}_{[TH]})'}{\partial \theta} \right] I_0^{-1} E \left[\frac{\partial m(\tilde{Y}_{[TH]})}{\partial \theta} \right] \right)^{-1}. \end{aligned}$$

In the above, $\frac{1}{T} \sum_{t=1}^T m(Y_t)$ is the vector of empirical moments based upon data Y_t of length T . When $m(Y_t) = Y_t$, for instance, we simply match the first moment of actual prices to the first moment of the simulated data. In general, the SMM matches several empirical moments based on the observed data with those simulated from the structural model under investigation. The simulations can be seen as approximating the population moments of the structural model by performing Monte Carlo integration, as determined by $\frac{1}{TH} \sum_{t=1}^{TH} m(\tilde{Y}_t)$. It is in this sense that the SMM is a simulation estimator. In a subsequent section, we will consider a special implementation of the estimator which bypasses the simulation step, and hence stops being ‘simulation based’. Nevertheless, because most empirical applications of SMM use simulations, we will continue to refer to the SMM as a ‘simulation estimator’.

In implementation, $(\partial m / \partial \theta)$ is obtained numerically by averaging over TH points of simulated data, and Ω^* is estimated as the long-run covariance matrix of $m(Y_t)$ using the Parzen window with four lags. The weights of the Parzen window are defined by:

$$\begin{aligned} w(x) &= \{1 - 6x^2 + 6|x|^3 \text{ if } 0 \leq |x| \leq 1/2 \\ &= \{2(1 - |x|)^3 \text{ if } 1/2 \leq |x| \leq 1. \end{aligned} \quad (4)$$

Note that Ω^* is estimated directly from the T sample points without any simulation.

For IND and EMM, we require an auxiliary model with parameters β of dimension q , where q must be greater or equal to the number of structural parameters (an identification condition). The auxiliary model is supposed to ‘adequately’ capture the properties of the observed data, though the model need not be correctly specified. There is no set rule in selecting an auxiliary model, and part of our analysis later in the paper is to provide a guide to finding the auxiliary model that suits the application in question. Let $Q(Y_{[T]}, \beta)$ be the objective function associated with the auxiliary model using actual data of length T , and let $Q(\tilde{Y}_{[TH]}, \beta)$ be the objective function for the same auxiliary model, but evaluated with the simulated data. Then $\hat{\beta}$ is estimated by maximizing the objective function $Q(Y_{[T]}, \beta)$ and $\tilde{\beta}$ is the maximand of the objective function $Q(\tilde{Y}_{[TH]}, \beta)$.

The IND begins with the premise that the auxiliary model is misspecified, and the simulations are meant to correct for the bias in the auxiliary model estimates induced by model misspecification. This is achieved by adjusting θ (the parameters of the structural model) such that $\hat{\beta}$ estimated with the observed data matches the $\tilde{\beta}$ estimated using the simulated data. The specifics are as follows:

IND

$$\begin{aligned}
D &= \hat{\beta}_T(Y_{[T]}) - \tilde{\beta}_{TH}(\tilde{Y}_{[TH]}, \theta) \\
\Omega^* &= J_0 I_0^{-1} J_0 \\
I_0^{-1} &= \lim_{T \rightarrow \infty} \text{Var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial Q(Y_t, \beta)}{\partial \beta} \right)^{-1} \\
J_0 &= -\frac{1}{T} \sum_{t=1}^T \frac{\partial^2 Q(Y_t, \beta)}{(\partial \beta \partial \beta')} \\
W_H^* &= \left(1 + \frac{1}{H} \right) \left(E \left[\frac{\partial^2 Q(\tilde{Y}_{[TH]}, \beta)}{\partial \theta \partial \beta'} \right] I_0^{-1} E \left[\frac{\partial^2 Q(\tilde{Y}_{[TH]}, \beta)}{\partial \beta \partial \theta'} \right] \right)^{-1} \\
\text{where } \hat{\beta} &= \text{argmax}_{\beta} Q(Y_{[T]}, \beta) \\
\text{and } \tilde{\beta} &= \text{argmax}_{\beta} Q(\tilde{Y}_{[TH]}, \beta).
\end{aligned}$$

The sample analogs of Ω^* and W_H^* are based on Appendix 2 of Gouriou, Monfort and Renault (1993). In particular, $\partial^2 Q / \partial \beta \partial \theta$ is obtained by taking the average (over the simulated paths) of the numerical derivatives of $\partial Q / \partial \beta'$ with respect to θ evaluated at $\hat{\theta}$. The quantity I_0 is approximated as the long-run variance of $\partial Q / \partial \beta$ in the observed data, once again using the Parzen window as the weighting function. It is interesting to note that Gouriou, Monfort and Renault (1993) used only an identity matrix as the weighting matrix in their empirical examples, on the ground that the efficiency gains from using the optimal weighting matrix are small. However, we feel it is desirable to put all three estimators on equal footing, and hence are prompted to use the optimal weighting

matrix to evaluate the IND.

The EMM estimator chooses θ such that the expectation of the gradient of the log density (the scores) for the auxiliary model evaluated with the density of the structural model, is close to zero. The expectation of the score is denoted as D below. For a simulated path of length N , we have:

EMM

$$\begin{aligned} D &= E_Y \left[\frac{\partial Q(Y_{[\infty]}, \beta)}{\partial \beta'} \right]_{\beta=\hat{\beta}} \approx \frac{1}{N} \sum_{n=1}^N \frac{\partial Q(\tilde{Y}_n, \hat{\beta})}{\partial \beta'}, \\ \Omega^* &= I_0^{-1} = \text{Var} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \frac{\partial Q(Y_t, \beta)}{\partial \beta} \right)^{-1}, \\ W^* &= \left(E \frac{\partial D(\tilde{Y}_{[\infty]})}{\partial \theta'} I_0^{-1} E \frac{\partial D(\tilde{Y}_{[\infty]})}{\partial \theta} \right)^{-1}, \end{aligned}$$

where $\partial D/\partial \theta$ is obtained by taking numerical derivatives, and I_0 is the long run variance of the scores using a four-lag Parzen window as weights. As in the other estimators, the optimal weighting matrix is directly computed from the sample data. However, the expectation of the scores can be calculated using other methods (such as quadrature or exploiting the probability distribution implied by the theory, for example). Therefore, the EMM method is not, strictly speaking, simulation based. It only becomes a simulation estimator when the expectation of the scores is approximated as the average over simulated data. This is the approach adopted in most empirical applications, and hence for a slight abuse of language, we also refer to EMM as a simulation based estimator. The EMM estimator is asymptotically as efficient as the MLE if the auxiliary model spans, or ‘smoothly embeds’ in the terminology of Gallant and Tauchen, the structural model.

A global specification test for the structural model can be obtained as

$$TH^* \times D' \Omega D \rightarrow \chi^2 \tag{5}$$

with $q - p$ degrees of freedom. For IND and SMM, $H^* = H/(1 + H)$. As can be seen from the definitions of W^* , increasing the number of simulations improves the efficiency of the SMM and IND estimates. For EMM, the length of the simulated series, N , is assumed to be ∞ . Since $T \cdot H = N$ by definition, $N \rightarrow \infty$ when $H \rightarrow \infty$. Hence $H^* = H/(1 + H) \rightarrow 1$ under EMM.

For identification purposes, IND and EMM require that the number of parameters in the auxiliary model be at least as large as the number of parameters in the structural model. For the SMM, we require that the number of moment conditions exceeds the number of parameters to be estimated, which is an order condition for identification. However, Duffie and Singleton (1993) note that the selected moment conditions should capture differing aspects of the data for identification of the structural parameters, though there is no formal way to check for this prior to estimation.

From a computational point of view, there are some fundamental differences between the three estimators. With the SMM, the practitioner only needs to specify the empirical moments and is the easiest to implement. The moments to be matched in the EMM are the scores, and hence the objective function Q is the log likelihood. Implementation of the EMM estimator is reasonably straightforward when the scores have simple forms. Since an infinite number of points has to be simulated from the structural model, at least in theory, the execution time of the EMM is largely determined by how difficult it is to simulate data from the structural model. In contrast, the number of replications in IND and SMM can be a much smaller number because the asymptotic variance-covariance matrix explicitly corrects for the simulation error by taking into account the finiteness of H . On the other hand, while the EMM method estimates the auxiliary model once, IND has to re-estimate it until convergence is achieved. This can be computationally burdensome when the auxiliary model is difficult to estimate.

Existing applications of the EMM estimator have primarily followed Gallant and Tauchen's suggestion of using the scores associated with the Semi-NonParametric density (SNP).⁴ We have not taken this route for three reasons. First, as we will see, the non-linearity in the conditional mean and variance of commodity prices is not of the type that can be captured by the SNP density. Second, examining the properties of the EMM estimator outside of the SNP environment is an interesting exercise in its own right. Third, we would like to compare the properties of EMM and IND using the same auxiliary model. But because we are running a monte-carlo, we need to take into account that use of complex auxiliary models for the IND can be time consuming. Thus, the selected auxiliary models all have a reasonably simple structure.

3 A Simple Example

We begin with an illustration of how the three estimators work in a very simple example. Suppose the true data generating process for a series $\{y_t\}_{t=1}^T$ is given by the following $MA(1)$ model:

$$y_t = e_t - \theta e_{t-1}. \tag{6}$$

The econometrician knows that the structural model is an $MA(1)$ and is interested in estimating the parameter θ using the simulation estimators. Our Monte Carlo is set up to assess the finite sample properties of $\hat{\theta}$. For each replication of the Monte Carlo experiment, a series $\{y_t\}_{t=1}^T$ is first obtained by evaluating (6) at the true value of θ and under the assumption that $e_t \sim N(0, 1)$ using a seed, say $seed_1$. The three simulation estimators then use (6) to simulate paths of $\{\tilde{y}_t\}_{t=1}^T$ for some initial guess of θ , also assuming $e_t \sim N(0, 1)$ but using a seed, $seed_2$, which differs from $seed_1$.

⁴See Gallant, Hsieh and Tauchen (1994) for a finance application using the SNP auxiliary model.

Series of $\{\tilde{y}_t\}_{t=1}^{TH}$ are re-simulated with updated values of θ (but the same $seed_2$) until a convergence criterion is satisfied.

For SMM, we consider the first two central moments, plus the first to third order autocovariances for a total of five empirical moments to be matched. That is,

$$m(y_t) = \left[y_t, (y_t - \bar{y})^2, (y_t - \bar{y})(y_{t-j} - \bar{y}) \right]', \quad j = 1, 2, 3. \quad (7)$$

whre \bar{y} is the sample mean.

For IND and EMM, we consider as auxiliary model an AR(3) to be maximized by maximum likelihood (conditional on y_1, y_2, y_3). Thus,

$$Q = -\frac{(T-3)}{2} \ln \sigma^2 - \sum_{t=4}^T \frac{(y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 y_{t-2} - \beta_3 y_{t-3})^2}{2\sigma^2}. \quad (8)$$

We first conduct a Monte Carlo experiment with $\theta = 0.5$ and $T=100$. For SMM and IND, we set $H = 10$. For EMM, $N = 1000$, so that all three models simulate the same number of data points. In the simulations, the `rndn()` function in Gauss is used to draw random numbers with $seed_1$ and $seed_2$ set at 255 and 2555 respectively, while the starting value of θ is set equal to 0.3.

As the results in Table 1 indicate, the estimators tend to under-estimate θ when $T = 100$. The IND is the most accurate both in terms of mean and median biases, with EMM coming in second, and SMM third. In terms of efficiency, SMM has the largest asymptotic standard errors, and IND the smallest. The standard deviation of the estimates in the simulations is reasonably close to the asymptotic standard errors. As far as the χ^2 statistic is concerned, the SMM has a size of 8.8% and is closest to the nominal size of 5%. However, the statistics associated with IND and EMM are both oversized (10.6% for IND, and 16.8% for EMM).

Increasing the sample size to 200 (see Table 1) improves the efficiency of all three estimators, and reduces the mean and median biases in EMM and SMM. Interestingly, the bias in IND is larger than when $T = 100$. Size distortions in the test for overidentification are also reduced when $T = 200$. The exact size of the test implied by SMM, IND and EMM is 5.6%, 7% and 11.8% respectively. Because the size distortion associated with EMM is still non-negligible, we increase the sample size and the number of simulations further to determine at what T and N the asymptotic approximation becomes adequate for the finite sample distribution of the test. Keeping $T=100$ and raising N to 5000 gives a mean estimate of .492 with an asymptotic standard error of .102 and a size of 15.6%. For $N=10000$, the mean is .491, the standard error is .102, and the size is 17.2%. Thus, raising N reduces the bias until $N=5000$, but does not help reduce the size distortions in the OI test. On the other hand, if we keep $N=10000$ but increase the sample size from $T=100$ to $T=500$, the mean estimate equals .498 with a standard error of .049 and a size equal to 8.6%. When the sample

size is increased further to $T=1000$, the mean estimate is .499 with an asymptotic standard error of .035, and the χ^2 test has a size of 6%. This suggests that the crucial improvement comes from increasing T rather than N and that a sample of 100 will not produce a very accurate test of over-identification restrictions when using the EMM.

A general overview of the properties of the estimators is given by the density plots of the estimates. The densities, along with a vertical line denoting the true value of θ , are given in Figure 1 for $T = 100$ and Figure 2 for $T = 200$. Figure 1 shows that the mode of the SMM estimates is slightly displaced from the true value of 0.5 and accounts for the larger mean and median biases reported in Table 1. Figure 2 shows that when T is increased to 200, all three estimators are centered around the true value, but the dispersion in the estimates is larger with SMM than with EMM and IND, hence accounting for the larger standard errors reported for SMM.

It is noteworthy to point out that the moments in this example are chosen to suit the structural model since we know that the MA(1) can be well approximated by a finite order AR model. An AR(3) auxiliary model was used for EMM and IND, for instance, while three of the moments used in SMM are meant to capture the autoregressive nature of the data generating process. In a more naive set of experiments, we purposely chose to match the first four central moments for SMM, and used a third order polynomial in y_{t-1} as the auxiliary model for EMM and IND. These moments would have been useful if the data had exhibited non-linearity, but this is not a feature of the MA(1) process. The estimates based on these moments were less efficient and exhibited higher mean and median biases. The choice of moments/auxiliary models, therefore matters, and the same qualitative results will bear out in the concrete economic example to follow.

4 The Speculative Storage Model

The economic environment with which the simulation estimators are to be tested is the speculative storage model under rational expectations with a non-negativity constraint on inventories. The model was first discussed in Williams (1937) and Samuelson (1957), and formalized in Gustafson (1958). Issues pertaining to estimating and testing of the model are discussed in Deaton and Laroque (1992, 1995, 1996). Here we only provide a brief description of the model.

Suppose the demand for commodities is given by the linear demand function:

$$p_t = P(z_t) = a + bz_t \tag{9}$$

where $a > 0$ and $b < 0$ are fixed parameters, $P(\cdot)$ is the inverse demand function, and z is an exogenous stochastic harvest process that is *i.i.d.* by assumption. Given this specification about linear demand, market clearing implies that the equilibrium price will follow the same process as the

harvest, since it is the sole source of supply. The problem becomes more interesting, yet complex, once speculators are introduced in the model. Inventory holders (speculators) are assumed to know the current harvest and demand the commodity in order to transfer it into the next period. They will do so whenever they expect to make a profit over storage and interest costs. Storing I units of the commodity leaves speculators with $(1 - \delta)I$ at the beginning of next period because of proportional deterioration. Let the discount rate be the constant, exogenous, real interest r . Arbitrage implies that

$$\begin{aligned} I_t &> 0 && \text{if } (1 - \delta)/(1 + r)E_t p_{t+1} = p_t, \\ I_t &= 0 && \text{otherwise.} \end{aligned} \tag{10}$$

The condition $I_t \geq 0, \forall t$ is the crucial source of non-linearity in the model. It forbids the market from borrowing commodities that have not yet been grown. Define x_t , the ‘amount on hand’, as the sum of the harvest and inherited inventories, so that

$$x_t = (1 - \delta)I_{t-1} + z_t. \tag{11}$$

In place of (9) we now have

$$p_t = \max \left[P(x_t), \frac{(1 - \delta)}{(1 + r)} E_t p_{t+1} \right]. \tag{12}$$

The first term would hold if the speculators believed t was the best time to sell and did not want to carry over inventories into the next period. The second term in the brackets is an intertemporal Euler equation that would hold provided $I_t > 0$, which in turn happens if the speculators expect to cover their discounted (after depreciation) costs from buying the commodity at time t .⁵

The autoregression function of prices is given by

$$E(p_{t+1}|p_t) = \min(p_t, p^*)(1 + r)/(1 - \delta). \tag{13}$$

An implication of (13) is that the overall price process is a non-linear first-order Markov process with a kink at $p^* = (1 - \delta)/(1 + r)E_t p_{t+1}$. Under the assumption that shocks to harvests are *i.i.d.*, p^* is a constant. It can also be shown that $\partial V(p_{t+1}|p_t)/\partial p_t \geq 0$.⁶ Intuitively, an increase in the current price reduces the incentive to hold inventories and therefore increases the chance of a stockout, which in turn increases the volatility of prices next period. The conditional variance increases in p_t until $p_t = p^*$ and is constant for $p_t > p^*$. Both the conditional mean and conditional variance, therefore, have kinks at p^* .

⁵The Euler equation is familiar from the consumption literature on liquidity constraints where the linear demand is replaced by a marginal utility function, see Deaton (1991).

⁶See Deaton-Laroque (1992).

Provided certain regularity conditions are met, in particular, $r + \delta > 0$ and z has a compact support, there exists a solution to (12) which defines the price function:

$$p_t = f(x). \quad (14)$$

Existence of a rational expectations equilibrium was proved in Deaton and Laroque (1992) under the assumption that the harvest is i.i.d.. The authors also showed that p_t is ergodic and hence has an invariant distribution. The functional equation that determines (14) is:

$$f(x) = \max \left\{ \frac{1-\delta}{1+r} \int f(z' + (1-\delta)(x - P^{-1}(f(x)))) d\Lambda(z'), P(x) \right\}, \quad (15)$$

where Λ is the distribution of the stochastic harvest process.

Various numerical techniques can be used to solve for $f(x)$. In this paper, we follow the solution method discussed in Deaton and Laroque (1992,1995,1996). In brief, this requires discretizing z and approximating Λ by a $M = 10$ point Gaussian distribution. The discretization procedure allows us now to write (15) as:

$$f(x) = \max \left[\frac{1-\delta}{1+r} M^{-1} \sum_{j=1}^M f \left(Z_j + (1-\delta)(x - P^{-1}[f(x)]) \right), P(x) \right]. \quad (16)$$

To calculate these price functions, we discretize the continuous state variable x in an equally spaced grid of gg points. Let X denote the discretized support of the continuous state variable x . In this notation, $X = X_1, \dots, X_{gg}$. A higher gg reduces approximation error but increases computational time. We used $gg = 50$ in this paper, and found no significant changes in the resulting price function or estimation results when gg is increased further. We use cubic splines to interpolate between grid points for values of x not on the grid. Since z is i.i.d., we have natural limits on the range which x can take. If the harvest is forever at its maximum level, and nothing is consumed out of it, then x asymptotes to Z_{max}/δ . At the other extreme, since inventories cannot be negative, the smallest value of x is the minimum harvest Z_{min} .

The basic procedure for solving (16) is as follows. Starting with $f_0(x) = \max(P(x), 0)$, we update the function using

$$f_{n+1}(x) = \max \left[\frac{1-\delta}{1+r} M^{-1} \sum_{j=1}^M f_n \left(Z_j + (1-\delta)(x - P^{-1}[f_n(x)]) \right), P(x) \right] \quad (17)$$

until $f_{n+1} = f_n$. Once $f(x)$ is found, and using the fact that $f(x)$ is monotone decreasing, we obtain $x_t = f^{-1}(p)$. Since

$$x_{t+1} = (1-\delta)[x_t - P^{-1}(f(x))] + z_{t+1}, \quad (18)$$

we have an equation for the evolution of prices:

$$p_{t+1} = f((1 - \delta)[f^{-1}(p_t) - P^{-1}(p_t)] + z_{t+1}). \quad (19)$$

The one caution with this solution algorithm is that there is no guarantee that $f^{-1}(p)$ lies in the range $[X_1, X_{gg}]$, where X_1 and X_{gg} are the lower and upper bounds of X , respectively. When too many points fall off the grid, the algorithm is stopped.

The structural parameters of this model are a, b, δ, r , the mean and the variance of the harvest. Given the identification proposition from Deaton-Laroque (1996), the mean and variance of the harvest cannot be separately identified from the linear demand parameters. Accordingly, the mean and variance are normalized to 0 and 1 respectively. This leaves four structural parameters to be estimated. Since we have limited interest in estimating r , the interest rate is fixed at .05, so that $\theta = (a, b, \delta)'$ are the parameters to be estimated. Furthermore, to ensure that $b < 0$ and $r + \delta > 0$, we estimate $b' = -exp(b)$, and $\delta' = -0.05 + exp(\delta)$, and recover the standard errors of b and δ using the delta method. The technical difficulty in estimating the model is that we have a bivariate stochastic process, comprised of x_t and z_t , both of which are unobserved, and are associated with an observable price through a non-linear equilibrium price function.

4.1 Finite Sample Properties of the PMLE

Deaton and Laroque (1996) estimated the commodity price model using the PMLE (pseudo maximum likelihood). The PMLE maximizes

$$2 \ln L = \sum_{t=1}^{t=T-1} \ln l_t = -(T-1) \ln(2\pi) - \sum_{t=1}^{t=T-1} \left(\ln s(p_t) - \frac{(p_{t+1} - \mu(p_t))^2}{s(p_t)} \right), \quad (20)$$

where $\mu(p_t) = E(p_{t+1}|p_t) \approx M^{-1} \sum_{j=1}^M f[(1 - \delta)(f^{-1}(p_t) - P^{-1}(p_t)) + z_j]$. The quantity $s(p_t) = V(p_{t+1}|p_t)$ is evaluated analogously (see Deaton and Laroque (1995) for details).

Results for the PMLE from a Monte Carlo experiment of 500 replications are given in Table 2⁷, and Figures 3 and 4 graph the densities of the three estimates. Although the finite sample distributions of the estimates resemble the normal distribution, the estimates are not centered around the true values. The bias is non-negligible and remains so as the sample size increases. As Deaton and Laroque (1995) noted, this could be because the PMLE makes no use of the theoretical predictions of the model about conditional skewness and kurtosis.

Before turning to the results for the simulation estimators, a note on non-convergent estimates is in order. Although in practice one can fine tune the initial guesses when convergence is not

⁷The true values for θ are set at (.6, -.3, .1) and the starting values at (.58, -.27, .12) as in the Monte-Carlo experiments done later on for the simulation estimators.

achievable, this is difficult to implement in a Monte Carlo experiment. Throughout our analysis, we discard the replication when we encounter non-convergence and draw a new sample. Our experience is that the PMLE produces far more non-convergent estimates than the simulation estimators to be considered. This could be an important issue when a more complex model is to be estimated. As Deaton and Laroque (1996) noted, the estimator performs poorly when the *i.i.d.* assumption is relaxed to allow for serially correlated harvests. It is with these properties of the PMLE in mind that we will be assessing the results for the simulation estimators.

5 Design of the Monte Carlo Experiment

In the terminology used when discussing the simulation estimators, $Y_t = p_t$. For each of the three simulation estimators, the speculative storage model is used as the structural model to simulate data. The dimension of θ is three.

To assess the three simulation estimators, we first use the following for SMM:

$$\boxed{\text{m1}} \quad m(p_t) = \left[p_t, (p_t - \bar{p})^2, (p_t - \bar{p})(p_{t-j} - \bar{p}) \right]', j = 1, 2, 3. \quad (21)$$

Summing these over T gives five sample moments. These moment conditions are somewhat naive as they do not explicitly recognize the non-linear nature of the price process. This is deliberate as we want to assess the sensitivity of $\hat{\theta}$ to the choice of moments, and hence the economic content required in the simulation stage for the simulation estimators to have good properties.

The second set of moment conditions incorporates more relevant information about the structural model by taking into account the skewness and kurtosis of the price process:

$$\boxed{\text{m2}} : \quad m(p_t) = \left[p_t, (p_t - \bar{p})^i, (p_t - \bar{p})(p_{t-1} - \bar{p}) \right]', \quad i = 2, 3, 4. \quad (22)$$

The moments defined by m_2 therefore allow for the non-linearity and serial dependence of commodity prices.

For IND and EMM, we consider four auxiliary models, and they are of increasing complexity. The first is an AR(3) with the log likelihood function:

$$\boxed{\text{M1}} : \quad Q_1 = -(T-3) \ln(2\pi) - \frac{(T-3)}{2} \ln \sigma^2 - \sum_{t=4}^T \frac{(p_t - \beta_0 - \beta_1 p_{t-1} - \beta_2 p_{t-2} - \beta_3 p_{t-3})^2}{2\sigma^2}. \quad (23)$$

The second auxiliary model is a third order polynomial in p_{t-1} :

$$\boxed{\text{M2}} : \quad Q_2 = -(T-1) \ln(2\pi) - \frac{(T-1)}{2} \ln \sigma^2 - \sum_{t=2}^T \frac{(p_t - \beta_0 - \beta_1 p_{t-1} - \beta_2 p_{t-1}^2 - \beta_3 p_{t-1}^3)^2}{2\sigma^2}, \quad (24)$$

again for a total of five parameters. Thus, in comparison to $M1$, non-linearity in the conditional mean is recognized, though the auxiliary model still assumes constant conditional variance.

The third auxiliary model allows for a time varying conditional variance. Its objective function is specified as:

$$\boxed{\text{M3}}: \quad Q_3 = -(T-1)\ln(2\pi) + \sum_{\tau=2}^T \left\{ -.5h_\tau - .5\frac{e_\tau^2}{\exp(h_\tau)} \right\}, \quad (25)$$

$$\begin{aligned} e_t &= p_t - \beta_0 - \beta_1 p_{t-1} - \beta_2 p_{t-1}^2, \\ \ln \sigma_t^2 &= h_t = \alpha_0 + \alpha_1 p_{t-1}^2. \end{aligned} \quad (26)$$

Model M3 allows for conditional heteroskedasticity, but does not capture the two-regime nature of the price process.

The final auxiliary model we considered is the Smooth Transition Autoregressive Model [STAR, see Granger and Terasvirtra (1993)] analyzed in Ng(1996) for the commodity price model under investigation. The model is defined as

$$p_t = \alpha_1 p^* - \frac{(\alpha_1 p^* - \alpha_1 p_{t-1})}{\Delta_t} + e_t \quad (27)$$

$$\begin{aligned} \ln \sigma_t^2 &= h_t = \alpha_0 + \frac{\beta_0 - \alpha_0 + \beta_1 p_{t-1}^2}{\Delta_t} \\ \Delta_t &= 1 + \exp(\gamma(p_{t-1} - p^*)) \end{aligned} \quad (28)$$

where γ is a large positive constant that controls how fast the model switches from one regime to the next. It has the property that if $p_{t-1} < p^*$, then $\Delta_t \rightarrow 1$, and thus p_t is an autoregression with conditional heteroscedasticity. If on the other hand, $p_{t-1} > p^*$, then $\Delta_t \rightarrow \infty$ and p_t is history independent with a fixed log variance. The parameters of the auxiliary model to be estimated are: $\beta = \{\alpha_1, \alpha_0, \beta_0, \beta_1, p^*\}'$, with γ fixed at (10/standard deviation of prices). Using the Kalman Filter to build up the likelihood function, the objective function for the auxiliary model is:

$$\boxed{\text{M4}}: \quad Q_4 = -(T-1)\ln(2\pi) + \sum_{\tau=2}^T \left\{ -.5h_\tau - .5\frac{e_\tau^2}{\exp(h_\tau)} \right\} \quad (29)$$

Model 4 is rich in economic structure as it explicitly models the two regime nature of the conditional mean and variance.

Given that the conditional ML estimates for M1 and M2 are the least squares estimates, estimation by maximum likelihood is not necessary. The estimates of M3 and M4 are, however, obtained by maximum likelihood.

Our Monte Carlo experiment consists of 500 simulations using Gauss 3.2.1. For the j^{th} Monte Carlo,

1. let $\theta = (0.60, -0.30, 0.10)'$ and compute the equilibrium price function;
2. simulate $\{e_t\}_{t=1}^T$ standard normal variates using seed 2555+j;
3. simulate $10 \cdot T$ points of p_t using (19). The last T points are treated as the data;
4. compute $m1$ or $m2$ in the case of SMM, and estimate auxiliary model $M1$, $M2$, $M3$, or $M4$ in the case of EMM and IND;
5. simulate standard normal variates $\{\tilde{e}_t\}_{t=1}^{1.1 \cdot N}$ using seed 255+j, where $TH \equiv N$;
6. start at the initial guess $\theta^0 = (0.58, -0.27, 0.12)$;
7. begin with iteration $i = 1$,
 - (a) for a given θ^i , solve for the equilibrium price function and use (19) to simulate $1.1 \cdot N$ points of \tilde{p}_t using $\{\tilde{e}_t\}$ as harvests; the last N points are retained.
 - (b) evaluate the criterion function $D'\Omega^*D$ and stop if convergence is achieved;
 - (c) update θ^i and return to (7a) (for the same j) until $D'\Omega^*D$ is minimized.
8. update j and go to 2.

As a computational matter, $\{e_t\}$ and $\{\tilde{e}_t\}$ can take on one of ten values because the Gaussian distribution is approximated by a ten point distribution (see Tauchen (1986)). In practice, we first simulate a Uniform(0,1) variable and use it as an indicator to select one of the ten points that approximate the Gaussian distribution. Parameter estimates are obtained using a form of the Berndt, Hall, Hall and Hausman (1974) algorithm, with numerical derivatives evaluated by small parameter perturbations. We use the gradient of the objective function as the convergence criterion, and the calculations are stopped when the elasticity of the criterion with respect to each parameter is less than 1.0E-4.

In addition to having four auxiliary models in the Monte Carlo, we let $T=100, 200$ and $N=500, 1000, 2500$. Thus, there are 24 cases in total. Note that although $f(X)$ is solved for a discrete number of points, each simulated data point satisfies the equilibrium price function. This is achieved by cubic spline interpolation between grid points. A particular experiment might fail for one of the following reasons: i) there are too many points in $f^{-1}(p)$ that fall off the grid; ii) the maximum likelihood used to estimate the auxiliary model fails to converge; or iii) the gradient of the objective function is flat and the iteration fails to converge. In such cases, the replication is discarded. The calculations are done on a Pentium Pro 200, a Pentium 166, a Pentium 120 with occasional help

from two unix workstations. It has proven computationally too demanding to assess the IND using M3 and M4 as the auxiliary models. Results for IND using M3 are restricted to 100 replications, and the attempt to obtain estimates using M4 is abandoned. We will elaborate on this issue below.

6 Results

The results for SMM are reported in Table 3. The SMM estimator tends to underestimate a and b and slightly over-estimates δ . For moments defined by $m1$, the estimates exhibit non-negligible mean and median biases especially when $T=100$. Indeed, these are some of the largest biases that we observe in the entire experiment. However, the estimates based on moments defined by $m2$ are much improved. Recall that the moments defined by $m1$ ignore the non-linearity in the data, and therefore do not capture as fully the data characteristics as $m2$. As a consequence, we have larger mean and median biases, by as much as around 15% for each of the three parameters when we use the $m1$ moments. Moreover, the estimates based upon $m1$ are less efficient, with standard errors for a , b and δ that are sometimes twice as large as those obtained by matching $m2$. While these results are not surprising, they stress the importance of which moments we choose to match. The moments should have enough variations to allow for identification of the structural parameters, as Duffie and Singleton (1993) emphasized. Ignoring the higher order moments (such as skewness and kurtosis) when the actual process is non-linear, for instance, will lead to both higher mean and median biases and less efficient estimates.

The asymptotic standard errors are derived under the assumption that $T \rightarrow \infty$. It is interesting to compare these with the standard deviation of the estimates in the Monte Carlo experiment, which can be thought of as the variation in the estimates when the number of replications becomes large (in this case, 500). While the asymptotic standard errors for a , b and δ are generally close to the standard deviation of the estimates, the discrepancies for SMM are much larger than what we record for IND and EMM (to be reported). For the $m1, T = 100$ combination, the asymptotic standard errors are sometimes less than half the standard deviations. Because such discrepancies are not observed for the simple MA(1) model reported in Table 1, one possible explanation is that the SMM estimates are more variable in practice than what is predicted by theory, as the structural model becomes increasingly complex.

The results for SMM also show that increasing TH from 500 to 1000 and 2500 did little to improve the bias and efficiency of the estimates. The estimates are improved on both counts, however, as T increases from 100 to 200. That is to say, increasing the length of the observed data has a stronger influence on the estimates than increasing the length of the simulated series. As for the test of over-identifying restrictions, the χ^2 test is slightly undersized for all variants of the

estimator considered.

The results for EMM for the four auxiliary models are reported in Table 4. The estimates are generally satisfactory, with small mean and median biases, even with the small number of simulated data we have used. While b is always under-estimated (in absolute terms), δ tends to be over-estimated. While the two linear models (M1 and M2) slightly over-estimate a on average, the two non-linear models slightly under-estimate it. Increasing T tends to reduce mean and median biases in the EMM estimator. Moreover, doubling the sample size reduces the standard errors of the estimates by the expected factor of $\sqrt{2}$.

The number of simulated data points, N , is much smaller than recommended in Gallant and Tauchen and warrants some investigation. In empirical finance applications, Gallant and Tauchen used an N of 100,000 when T is 16,127. Because we only have 88 observations for commodity prices, the N used in financial applications may not be appropriate for our application. Since $N = T \cdot H$, we essentially experimented with simulating 5 to 25 times the length of the observed data. We have found only a slight bias reduction as N increases from 1000 to 2500. Increasing N reduces the asymptotic standard errors of the estimates as one would expect. While it is possible that we have put the EMM on unfair ground because N is too small, in a subsequent section, we will present results for EMM which amount to letting $N \rightarrow \infty$.

While the estimates using EMM are similar across auxiliary models, there are two notable differences. The first concerns the OID test. The two linear models have rejection rates of zero, while the two non-linear models have reasonable rejection rates, yet with variations that seem unrelated to either N or T . This finding suggests that the choice of auxiliary model matters. In our controlled setting, we know that prices are non-linear in mean and variance, and auxiliary models which capture these characteristics apparently have better size properties. The second difference arises in the standard deviation of the estimates relative to the asymptotic standard errors. The discrepancies in the two statistics are generally larger for the two linear models. In particular, the variations in $\hat{\delta}$ are sometimes twice as large as those suggested by the asymptotic standard errors. Although discrepancies between the two statistics are reduced as N and T increase, they remain non-negligible for the two linear auxiliary models. A similar problem was observed for the SMM with the auxiliary models which do not capture the nonlinearity in the data. This suggests that the asymptotic standard errors of the SMM and EMM may understate the variability of the estimates when the auxiliary model does not capture ‘well’ the features of the actual prices.

Results for IND are reported in Table 5. We first focus on the two linear models. The estimates have small mean and median biases. The estimator tends to under-estimate a and b , and over-estimates δ . Unlike the EMM estimator, there appears to be less variation in the estimates across

auxiliary models. This might be due to the fact that the IND begins with the premise that the auxiliary model is not a correctly specified model, and therefore it matters less how well the auxiliary model fits the data. Accordingly, the estimates from both auxiliary models have similar properties. The asymptotic approximation of the standard errors is also reasonably good. The χ^2 test is consistently in the neighborhood of 5%. One interesting aspect of the IND is that the efficiency gain from increasing T and/or N is less noticeable than in SMM and EMM.

Our experience with applying the indirect inference estimator to non-linear auxiliary models has been discouraging. To be fair, we do not believe that the IND was intended to be used with complex auxiliary models, and evaluating the IND under M3 and M4 was more an issue of completeness. Nevertheless, we encounter two problems with the IND when non-linear auxiliary models are used. The first is the tremendous demand on time, because the auxiliary model has to be re-estimated until convergence is achieved. It also appears that convergence is more easily obtained by averaging over H estimates from simulated data of length T rather than simulating one path of length $T \cdot H$, even though the two methods should produce estimates with the same asymptotic properties. This route increases computational time even further, however, since the non-linear model has to be estimated H times per iteration. We have little choice but to restrict the number of Monte Carlos to 100 for M3. Second, in a lot of cases, we fail to obtain estimates that satisfy our convergence criterion. In fact, difficulties with using non-linear models surface even when working with the MA(1) example. For M4, the problem was severe enough that we have no choice but to abandon the experiment. In any event, we have not found using the non-linear auxiliary model to be worth the computation cost. Results for M3 in Table 5 suggest that the estimates are generally more biased and less efficient than the indirect estimates based upon the two linear models.

For each of the three estimators, the densities of the ‘best’ set of estimates for $T=100$ are graphed in Figure 5 and for $T=200$ in Figure 6. More precisely, these densities are based on $m2$ for SMM, $M4$ for EMM and $M2$ for IND. For the parameter a , EMM is centered closest to the true value and also has the smallest dispersion. For b , SMM and IND dominate EMM both in terms of location and dispersion. For δ , all three estimators produce outliers that account for the long tail in the distribution of the estimates. The possibility for large outliers is one undesirable finite sample property in all three estimators that our Monte Carlo experiments reveal.

In terms of the computational aspects of the three estimators, we have the following observations. First, in terms of the time necessary to “babysit” the Monte Carlo experiments, the SMM is by far the easiest. The simulations for a given configuration (of T for various values of N) take at most one day to compute. For EMM, it depends on N . For N no larger than 2500, as reported here, the simulations for a configuration take a little under two days. However, when N is increased

to 10,000, as in results not reported in the paper, the time required triples. This is because the Markov nature of the price process requires generating p_{t+1} from p_t sequentially, and this is a rather time consuming operation in Gauss when N is large. For IND, the time required strongly depends on the auxiliary model. For the two linear models, the time required is only slightly more than EMM. For the non-linear models, it takes several days just to do 100 Monte Carlos with $N = 500$. Thus, from a pure computational standpoint, SMM comes out the winner, followed by EMM, and then IND.

Compared with the PMLE estimator, the best estimates produced by the simulation estimators have a lower bias. However, because the simulation estimators produce outliers more frequently, they also tend to be less efficient. Thus, in terms of efficiency, the PMLE still holds an advantage over the simulation estimators, at least for the configurations of N and T being considered. In the next section, we will investigate the trade-off between bias and efficiency.

7 Using the Invariant Distribution

Deaton-Laroque (1992) noted that because of the non-linearity of the price function and the occasional spikes predicted by the model, a large number of simulated prices is required for an accurate evaluation of the moments. However, the speculative commodity storage model predicts that prices follow a renewal process and thus have an invariant distribution [see Deaton-Laroque (1992)]. The existence of an invariant distribution is convenient because for two of the estimators considered, the invariant distribution allows us to generate the (population) moments without having to approximate them by Monte Carlo simulations. In other words, use of the invariant distribution amounts to the limiting case where the number of simulated prices is infinite. The invariant distribution thus allows for rapid and accurate evaluation of the moments necessary to obtain the parameter estimates. Indeed, one of the advantages of using the commodity price model for the Monte Carlo experiment is that it allows us to examine the extent to which simulation biases can be reduced by increasing the number of simulations.

To calculate the invariant distribution associated with prices, define $w(x_t)$ implicitly from rewriting the transition equation (18):

$$x_{t+1} = (1 - \delta)(x_t - P^{-1}(f(x_t))) + z_{t+1} = w(x_t) + z_{t+1} \quad (30)$$

Recall that the compact support of the continuous state variable x is discretized by an equidistant set of grid points, X , where $X = (X_1, \dots, X_{gg})$ and $gg = 50$ in our computations. The transition probability from X_j to X_i is

$$R_{ij} = pr(X_i + \Delta/2 \geq x_{t+1} \geq X_i - \Delta/2 | x_t = X_j) \quad (31)$$

where Δ is the distance between points on the grid X . Substituting from (30) and using the assumption that z is a standard normal variate, we have

$$R_{ij} = \Phi(X_i + \frac{\Delta}{2} - w(X_j)) - \Phi(X_i - \frac{\Delta}{2} - w(X_j)) \quad (32)$$

Given the matrix R , the probabilities associated with the states are updated by

$$\pi_{it+1} = \sum_{j=1}^{gg} R_{ij} \pi_{jt}. \quad (33)$$

The invariant distribution π is then the normalized eigenvector of R corresponding to the unit eigenvalue. Thus, π can be obtained by solving the linear equations

$$\begin{pmatrix} R - I & e \\ e' & 0 \end{pmatrix} \begin{pmatrix} \pi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (34)$$

where e is an M -vector of ones. Therefore, for every point on the grid X_i , there is an associated probability π_i , and this is also the probability of observing $f(X_i)$. We now discuss how the invariant distribution can be exploited in implementing EMM and SMM.

7.1 EMM using the invariant distribution

The EMM matches the (unconditional) expectation of the scores of the auxiliary model, evaluated with the density of the structural model. For the commodity price model, the t^{th} element of a typical score is a function of p_t , p_{t-1} and $\hat{\beta}$. Denote this by $q(p_t, p_{t-1}, \hat{\beta})$. The EMM estimator implemented in the previous sections evaluates the expectation of q by Monte Carlo integration. That is, simulating the price process from the assumed structural model, and then averaging over the simulated data. However, the expectation of the scores can also be calculated by using the law of iterated expectations, in conjunction with the invariant distribution. Specifically, we can first integrate out terms involving p_{t-1} only, and because $p_t = f(x) \forall t$, p_{t-1} can be replaced by $f(X_i)$ in the discretized state space. For a given X_i , $E(p_t | p_{t-1})$ is given by

$$ep_i = M^{-1} \sum_{j=1}^M f[(1 - \delta)(X_i - P^{-1}(f(X_i))) + z_j], \quad (35)$$

in view of (19). The expectation of the scores is then calculated as a weighted sum of $q(ep_i, f(X_i), \hat{\beta})$, where the weights are the unconditional probabilities associated with the invariant distribution. An example makes this clear.

Consider the score with respect to the constant term β_0 in the M3 auxiliary model:

$$\frac{\hat{e}_t}{\exp(\hat{h}_t)} \quad (36)$$

where $\hat{e}_t = p_t - \hat{\beta}_0 - \hat{\beta}_1 p_{t-1} - \hat{\beta}_2 p_{t-1}^2$ and $\hat{h}_t \equiv \ln(\sigma^2) = \hat{\alpha}_0 + \hat{\alpha}_1 p_{t-1}^2$. The conditional expectation of (36), conditioned on information at time $(t-1)$, for a given X_i is

$$q_i = \frac{ep_i - \hat{\beta}_0 - \hat{\beta}_1 f(X_i) - \hat{\beta}_2 [f(X_i)]^2}{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 [f(X_i)]^2)} \quad (37)$$

where ep_i is given by (35). The expectation of the score is computed as

$$\sum_{i=1}^{gg} q_i \cdot \pi_i \quad (38)$$

Note, however, that the methodology is applicable to $M2, M3$ and $M4$, but not to $M1$ because of $M1$'s AR(3) structure. The law of iterated expectations cannot be used in this case, because the model does not predict any tractable relationship about $E(p_t | p_{t-3})$, for instance.

7.2 SMM using the invariant distribution

Using the invariant distribution to obtain the moments required by the SMM is more straightforward. To evaluate a particular moment, (say the first), we simply use

$$\sum_{i=1}^{gg} f(X_i) \cdot \pi_i - \frac{1}{T} \sum_{t=1}^T p_t, \quad (39)$$

where $f(X_i)$ is the equilibrium price function evaluated at $x = X_i$, and π_i is the unconditional probability associated with X_i in the discretized state space. The second, third and fourth moments are defined analogously. Note that for the first four central moments we do not need to resort to the law of iterated expectations (as was necessary in the case of EMM) because these moments are already in unconditional form. However, we do need to evaluate the first order autocovariance, and in that case, we once again appeal to the law of iterated expectations. More specifically, we need to evaluate

$$E_p[(p_t - \mu)(p_{t-1} - \mu)]. \quad (40)$$

Conditioning on p_{t-1} and then integrating out p_{t-1} , we have

$$\sum_{i=1}^{gg} (ep_i - \mu) \cdot (f(X_i) - \mu) \cdot \pi_i \quad (41)$$

where ep_i is defined as in (35), $f(X_i)$ is the equilibrium price function and $\mu = \sum_{i=1}^{gg} f(X_i) \cdot \pi_i$.

The invariant distribution can be used to construct the SMM estimates when used in conjunction with the moments defined by $m2$ but not $m1$. We cannot implement $m1$ because the model does not predict any tractable relationship about $E[p_t | p_{t-j}]$ for $j > 1$.

7.3 Estimation Based on the Invariant Distribution and Simulation Bias

SMM results based on the invariant distribution are given in Table 6. Surprisingly, the estimates have somewhat larger bias than those based on Monte Carlo integration reported in Table 3. Offsetting estimation and simulation errors could account for the smaller bias in the simulation based estimates. However, there is a noticeable gain in efficiency. In most cases, the standard errors based on the invariant distribution are reduced by as much as 50%.

The results for EMM based on the invariant distribution are also given in Table 6. The biases are smaller than those based on 1000 simulated prices using M4 when $T=100$, and are of similar magnitude when $T=200$ (recall that M4 is the auxiliary model that yields estimates with the lowest bias and highest efficiency (see Table 4)). However, with the less adequate auxiliary models such as M2 and M3, use of the invariant distribution either leads to a substantial reduction in bias (especially in the parameter b) or an improvement in efficiency, but rarely both. For example, the standard errors for b and δ are generally smaller when the invariant distribution is used, but the biases are larger. Overall, a larger number of simulations seems to partially compensate for the inadequacy of the auxiliary model. There is also reduced variability in the estimates when the invariant distribution is used; the standard deviations of the estimates are smaller. Note that for all auxiliary models considered, the large efficiency gains seem to occur as T , rather than N , increases. Indeed, at $T=200$, the results using the invariant distribution are only slightly better than those with $N = 2500$. Thus, as emphasized earlier when evaluating the MA(1) example, increasing T seems to have a stronger effect on the properties of EMM than increasing N , especially once N has reached a certain critical point. For the present application, this value of N is around 1000, or 10 times the sample size.

The average mean squared error (MSE) of the estimates based on the best auxiliary model and moments is presented in Table 7. For all three estimators, an increase in the sample size induces a larger reduction in MSE than an increase in the number of simulated data points. The EMM always produces the best estimates for the parameter a , but the worst estimates of b . On the other hand, the SMM produces the best estimates of b , but the worst estimates of a and δ . The IND comes second in all three parameters.

The difference in MSE between the simulation based and the invariant distribution based estimates can be interpreted as simulation error. For EMM, the simulation error induced on \hat{b} is the largest, but it falls rapidly with the sample size. At $T=200$, the MSE for all three parameters estimated by EMM is quite close to the MSE of the invariant distribution. The results once again suggest that for EMM, T is very important to the properties of the estimates, more so than the number of simulations. For the SMM, the simulation error in \hat{a} is the largest, but it falls as either

N or T rise.

It is useful to have a single metric to assess how N affects the simulation bias. For this purpose, we consider the mean squared error of prices. This is defined as the mean squared difference between the prices derived from the structural model evaluated at the true value of θ , with those evaluated at the estimated parameters. The results are reported in the last column of Table 7. The MSE for prices is the smallest for the invariant distribution and increases as N is reduced (for a given estimator). This was expected because the invariant distribution is equivalent to simulating an infinite number of prices. It is nonetheless surprising to find that for EMM using M4, setting N to ≈ 2500 when $T = 200$ does almost as well as using the invariant distribution (with MSE of .1176 and .1165 respectively). With $T = 100$, the decrease in simulation bias from increasing N beyond 1000 is small (the MSE is .3222 for $N = 1000$ and falls to .2914 when $N = 2500$). These results suggest that performing the Monte Carlo integration over $N \approx 10T$ observations is adequate. The SMM results support these conclusions. The decrease in MSE when raising N from 1000 to 2500 is small (.3588 to .3454 when $T = 100$), and it is not far off from the invariant distribution bias (.2850). Indeed, when Gallant and Tauchen simulated 100,000 data points, they are simulating roughly seven times the size of the observed data series. Our suggestion of $N \approx 10T$ is slightly on the conservative side.

Because no one estimator is superior in estimating all three parameters, the MSE of prices also provides a convenient metric for an overall comparison. The result that stands out is that all three simulation estimators are inferior to the PMLE. As mentioned earlier, the PMLE is mean and median biased. Although the simulation estimators provide improvements on this count, the estimates are also less variable and on net, have larger errors than the PMLE in a mean-squared sense. Of the three simulation estimators, the SMM is noticeably the weakest. For $T = 100$, the EMM with a large N edges out IND very slightly. The SMM and EMM exhibit sharper reductions in MSE than the IND as the sample size increases. For $T=200$, EMM has the smallest MSE even when a small N . It appears that sample size permitting, the EMM is the best of the three estimators from a mean-squared error point of view. For smaller sample sizes, the IND is a stable performer and it is the least sensitive to variations in N or T .

8 Conclusion

In this paper, we use the speculative storage model under rational expectations as the test bed for three estimators which are suited when the objective function for classical estimation becomes intractable. There are some tradeoffs among the estimators worthy of consideration.

1. Statistical efficiency vs. computational efficiency:

The SMM tends to have larger biases and variances but is the easiest to implement. The IND is more robust to variations in N and T but can be computationally costly when the auxiliary model is difficult to estimate. The EMM is best when N and especially T are large, but its computational cost increases with N and the complexity of the structural model.

2. Choice of moments and auxiliary models:

For the SMM and the EMM, choose moments and an auxiliary model to encompass as many features (such as non-linearity) in the data as possible. On this, economic theory can provide a guide. For IND, keep to a simple auxiliary model.

3. Number of simulated data points:

Use of the theoretical invariant distribution to evaluate the relevant expectations yields efficiency gains over Monte Carlo integration. When the expectations are to be approximated by simulations, simulating ten times the number of observations as in the actual data seems appropriate in the sense that the reductions in simulation bias from simulating more observations is small.

4. The sample size:

The properties of the EMM are the most sensitive to the sample size. At $T = 200$, the EMM comes closest to the efficiency of the PMLE, and the difference should narrow as T increases. Nonetheless, the EMM-based OID test has size distortions until T is fairly large. When conducting inference, this issue should be borne in mind.

Although none of the simulation estimators beat the PMLE in a mean-squared sense, they do have better bias properties, and convergence is generally less of an issue with the simulation estimators. Thus, in dynamic models of complex structure, the simulation estimators can be a useful complement, if not an alternative, to the PMLE.

References

- [1] Dixit Avinash and Pindyck Robert, 'Investment under Uncertainty' 1994 Princeton University Press.
- [2] Altug, S. 'Time to build and aggregate fluctuations: some new evidence', *International Economic Review*, 30, 1989, pp. 889-920.
- [3] Bansal R., Gallant R., Hussey R., and Tauchen G., 'Non parametric estimation of structural models fir high frequency currency market data' *Journal of Econometrics*, 66, 1995, pp.251-287.
- [4] Berndt, Ernst., Brownyn H. Hall, Robert E. Hall, and Jerry A. Hausman, 1974, 'Estimation and Inference in Nonlinear Structural Models,' *Annals of Economic and Social Measurement*, 3, 653-666.
- [5] Carroll, Christopher D., 1997, 'Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis' *Quarterly Journal of Economics*, CXII, 1, 1659-1686.
- [6] Carroll, Christopher and Andrew Samwick. 1997.'The Nature of Precautionary Wealth.' *Journal of Monetary Economics*, 40(1).
- [7] Carroll, Christopher and Andrew Samwick. 1998.'How important is precautionary saving?' *Review of Economics and Statistics* 80 no. 3: 410-19.
- [8] Deaton Angus, 'Saving and Liquidity Constraints',1991, *Econometrica*, 59, No.5, 1221-1248.
- [9] Deaton Angus and Laroque Guy, 'On the behavior of commodity prices' *Review of Economic Studies*, 1992, pp.1-23.
- [10] ————— 'Estimating a nonlinear rational expectations commodity price model with unobserved state variables' 1995, *Journal of Applied Econometrics*, S3-S55.
- [11] ————— 'Competitive Storage and Commodity Price Dynamics' *Journal of Political Economy*, 1996, vol.104, no.5, pp. 896-923.
- [12] Duffie Darrell and Singleton Kenneth. 'Simulated Moments Estimation of Markov Models of Asset Prices' *Econometrica* Vol.61, No. 4 July 1993, 929-952.
- [13] Gallant Ronald and Tauchen George,'Which Moments to Match' *Econometric Theory*, 12, 1996, 657-681.
- [14] Gallant Ronald, Hsieh David and Tauchen George (1994) 'Estimation of Stochastic Volatility Models with Diagnostics' Working Paper, Duke University.
- [15] Gourieroux C., Monfort A., and Renault. 'Indirect Inference' *Journal of Applied Econometrics*, vol. 8, S85-S118 (1993).
- [16] Gourinchas P. and Parker J., 'Consumption over the Life Cycle', 1996, MIT mimeo.
- [17] Gross David, 'The investment and financing decisions of liquidity constrained firms' 1994 MIT mimeo.
- [18] Gustafson R. L. 'Carryover Levels for Grains' (US Department of Agriculture, Technical Bulletin, 1178), 1958.

- [19] Hansen Lars, ‘Large Sample Properties of Method of Moments Estimators’ *Econometrica*, 1982, 50, 1029-1054.
- [20] Hotz, J. and R. Miller, 1993, ‘Conditional Choice Probabilities and the Estimation of Dynamic Programming Models’ *Review of Economic Studies* 60:497-530.
- [21] Ludvigson Sydney, ‘Consumption and credit: a model of time varying liquidity constraints’ *forthcoming, Review of Economics and Statistics*.
- [22] Ludvigson, Sydney and Alexander Michaelides 1999. ‘Can Buffer Stock Saving Explain the Consumption Excesses?’ Federal Reserve Bank of New York Working Paper.
- [23] McGrattan Ellen, Rogerson Richard, and Wright Randall, ‘An Equilibrium Model of the Business Cycle with Household Production and Fiscal Policy’ Federal Reserve Bank of Minneapolis, Staff Report 191.
- [24] Ng Serena, ‘Looking for evidence of speculative stockholding in commodity markets’ *Journal of Economics Dynamics and Control* 20 (1996) 123-43.
- [25] Ohanian, Lee, Giovanni Violante, Per Krusell, and José-Victor Ríos-Rull, ‘Simulation-Based Estimation of a Nonlinear Latent Factor Aggregate Production’, 1997, mimeo, Federal Reserve Bank of Minneapolis.
- [26] Pakes, A. (1994) ‘Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interactions’, in C. A. Sims (ed.), *Advances in Econometrics: Sixth World Congress*, 2, Cambridge University Press, Cambridge.
- [27] Rust, J. (1987) ‘Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher’ *Econometrica* 55: 999-1057.
- [28] Rust, J. (1994) ‘Estimation of dynamic structural models, problems and prospects: discrete decisions’, in C. A. Sims, (ed.), *Advances in Econometrics: Sixth World Congress*, 2, Cambridge University Press, Cambridge.
- [29] Samuelson P. A. ‘Stochastic Speculative Price’ 1971, Proceedings of the National Academy of Science, 68, 335-37.
- [30] Tauchen G., ‘Finite State Markov Chain Approximations to Univariate and Vector Autoregressions’, 1986, *Economics Letters*, 20, 177-181.

Table 1: MA(1)

Method	Moments	T	N	Mean	A.S.E.	OID
				Median	Std.	
GT	M1	100	1000	.480	.100	.168
				.480	.107	
IND	M1	100	1000	.498	.086	.106
				.498	.108	
SMM	m1	100	1000	.469	.130	.088
				.456	.140	
GT	M1	200	1000	.497	.075	.118
				.496	.071	
IND	M1	200	1000	.509	.070	.070
				.509	.073	
SMM	m1	200	1000	.495	.110	.056
				.493	.120	

N is the length of the simulated series and T is the length of the data. A.S.E. denotes asymptotic standard error, while std. denotes the standard deviation of the estimate in 500 replications. The OID is the test for overidentifying restrictions and has as a χ^2 distribution with $p - q$ degrees of freedom, where p is the number of structural parameters, and q is the number of moments in the case of SMM and the number of parameters of the auxiliary model in the case of IND and GT. The moments (models) are defined in the text.

Table 2: Results for PMLE

Model	T	N=TH	a		b		δ		OID
			Mean Median	A.S.E. std.	Mean Median	A.S.E. Std.	Mean Median	A.S.E. Std.	
PMLE	100	—	.605	.032	-.289	.040	.108	.029	—
			.604	.029	-.289	.037	.107	.026	
PMLE	200	—	.603	.022	-.292	.028	.106	.020	—
			.603	.021	-.289	.026	.106	.019	

Table 3: Results for SMM

Model	T	N=TH	a		b		δ		OID
			Mean Median	A.S.E. std.	Mean Median	A.S.E. Std.	Mean Median	A.S.E. Std.	
m1	100	500	.557	.047	-.279	.063	.123	.075	.032
			.562	.110	-.271	.055	.105	.073	
m1	100	1000	.560	.046	-.276	.059	.124	.069	.034
			.564	.110	-.270	.053	.124	.069	
m1	100	2500	.560	.044	-.276	.058	.124	.068	.034
			.564	.100	-.270	.050	.110	.070	
m1	200	500	.570	.033	-.290	.048	.110	.050	.032
			.570	.090	-.286	.048	.101	.048	
m1	200	1000	.580	.032	-.290	.043	.110	.050	.044
			.580	.082	-.287	.042	.102	.046	
m1	200	2500	.580	.030	-.290	.041	.110	.046	.058
			.580	.080	-.287	.040	.104	.043	
m2	100	500	.593	.082	-.289	.021	.110	.027	.018
			.593	.056	-.293	.036	.105	.048	
m2	100	1000	.597	.078	-.290	.020	.108	.025	.052
			.599	.052	-.294	.032	.101	.042	
m2	100	2500	.597	.076	-.288	.019	.109	.025	.02
			.597	.049	-.293	.033	.102	.044	
m2	200	500	.598	.036	-.298	.018	.101	.031	.024
			.596	.065	-.298	.016	.102	.019	
m2	200	1000	.600	.035	-.297	.017	.102	.030	.022
			.598	.060	-.298	.015	.102	.018	
m2	200	2500	.600	.033	-.297	.016	.101	.028	.024
			.599	.050	-.297	.014	.100	.017	

Table 4: Results for GT

Model	T	N=TH	a		b		δ		OID
			Mean Median	A.S.E. std.	Mean Median	A.S.E. Std.	Mean Median	A.S.E. Std.	
M1	100	500	.602	.030	-.290	.060	.120	.084	.000
			.603	.021	-.287	.120	.105	.200	
M1	100	1000	.602	.029	-.286	.058	.124	.082	.000
			.602	.021	-.281	.120	.106	.150	
M1	100	2500	.600	.028	-.287	.055	.123	.081	.000
			.600	.021	-.284	.120	.123	.081	
M1	200	500	.601	.020	-.296	.050	.120	.070	.000
			.601	.016	-.290	.087	.100	.100	
M1	200	1000	.600	.020	-.295	.051	.11	.060	.000
			.600	.015	-.290	.087	.100	.095	
M1	200	2500	.600	.020	-.294	.048	.120	.066	.000
			.600	.016	-.291	.087	.100	.099	
M2	100	500	.605	.036	-.278	.045	.116	.049	.000
			.605	.055	-.278	.045	.112	.100	
M2	100	1000	.600	.036	-.277	.042	.116	.047	.000
			.604	.055	-.273	.120	.112	.100	
M2	100	2500	.603	.034	-.276	.043	.117	.046	.000
			.604	.056	-.272	.120	.114	.100	
M2	200	500	.603	.028	-.285	.038	.110	.038	.000
			.603	.039	-.281	.093	.109	.070	
M2	200	1000	.603	.028	-.285	.035	.110	.036	.000
			.604	.041	-.280	.099	.108	.072	
M2	200	2500	.603	.026	-.284	.033	.112	.034	.000
			.603	.04	-.281	.093	.110	.072	
M3	100	500	.598	.030	-.276	.048	.110	.043	.044
			.598	.024	-.275	.033	.105	.031	
M3	100	1000	.598	.030	-.276	.044	.109	.040	.036
			.598	.024	-.273	.032	.105	.031	
M3	100	2500	.598	.029	-.274	.043	.109	.041	.030
			.598	.024	-.273	.032	.105	.031	
M3	200	500	.599	.022	-.289	.032	.104	.029	.060
			.599	.018	-.288	.024	.100	.021	
M3	200	1000	.600	.020	-.289	.030	.104	.028	.042
			.600	.018	-.289	.024	.102	.022	
M3	200	2500	.600	.020	-.288	.028	.105	.026	.054
			.600	.018	-.287	.024	.102	.022	
M4	100	500	.598	.038	-.288	.057	.100	.037	.058
			.598	.025	-.289	.040	.099	.025	
M4	100	1000	.598	.038	-.286	.056	.100	.034	.082
			.599	.025	-.287	.040	.099	.026	
M4	100	2500	.598	.036	-.286	.056	.101	.026	.078
			.599	.025	-.289	.039	.099	.026	
M4	200	500	.598	.025	-.296	.038	.099	.025	.112
			.597	.018	-.294	.029	.097	.018	
M4	200	1000	.599	.023	-.297	.035	.099	.024	.068
			.599	.019	-.295	.028	.099	.018	
M4	200	2500	.598	.023	-.296	.033	.099	.023	.076
			.597	.019	-.292	.028	.098	.019	

Table 5: Results for IND

Model	T	N=TH	a		b		δ		OID
			Mean Median	A.S.E. std.	Mean Median	A.S.E. Std.	Mean Median	A.S.E. Std.	
M1	100	500	.596	.025	-.285	.054	.130	.079	.082
			.595	.032	-.278	.056	.116	.080	
M1	100	1000	.596	.025	-.282	.055	.129	.065	.078
			.596	.03	-.276	.055	.120	.070	
M1	100	2500	.595	.025	-.280	.054	.127	.064	.084
			.596	.03	-.272	.052	.120	.066	
M1	200	500	.598	.020	-.294	.042	.120	.043	.066
			.599	.022	-.290	.046	.108	.053	
M1	200	1000	.597	.020	-.290	.041	.119	.043	.052
			.597	.020	-.288	.042	.110	.048	
M1	200	2500	.599	.020	-.291	.041	.112	.040	.074
			.596	.020	-.289	.036	.109	.043	
M2	100	500	.592	.034	-.298	.038	.103	.038	.086
			.593	.043	-.296	.045	.100	.036	
M2	100	1000	.590	.036	-.295	.039	.104	.030	.088
			.590	.040	-.295	.044	.105	.035	
M2	100	2500	.590	.036	-.292	.039	.104	.031	.086
			.592	.040	-.293	.043	.104	.035	
M2	200	500	.594	.027	-.308	.029	.099	.022	.074
			.597	.034	-.303	.037	.098	.029	
M2	200	1000	.594	.028	-.303	.029	.100	.022	.076
			.594	.032	-.300	.032	.100	.027	
M2	200	2500	.593	.028	-.301	.025	.100	.022	.088
			.594	.03	-.298	.030	.099	.026	
M3	100	200	.603	.019	-.322	.024	.167	.028	.02
			.603	.032	-.314	.058	.146	.108	

Table 6: Results for GT and SMM : The Invariant Distribution

Model	T	N=TH	a		b		δ		OID
			Mean Median	A.S.E. std.	Mean Median	A.S.E. Std.	Mean Median	A.S.E. Std.	
M2 GT	100	INV	.595	.025	-.269	.033	.104	.032	.126
			.594	.027	-.269	.046	.097	.045	
M2 GT	200	INV	.595	.025	-.283	.025	.099	.022	.130
			.594	.019	-.282	.029	.096	.024	
M3 GT	100	INV	.595	.027	-.269	.046	.104	.045	.126
			.594	.025	-.269	.033	.097	.032	
M3 GT	200	INV	.595	.02	-.284	.029	.099	.024	.130
			.594	.018	-.282	.025	.099	.024	
M4 GT	100	INV	.603	.032	-.296	.049	.097	.032	.078
			.604	.026	-.296	.040	.096	.027	
M4 GT	200	INV	.600	.023	-.297	.033	.099	.022	.062
			.599	.019	-.298	.030	.098	.019	
m2 SMM	100	INV	.596	.068	-.279	.018	.120	.022	.026
			.596	.044	-.282	.025	.116	.041	
m2 SMM	200	INV	.595	.050	-.285	.013	.112	.015	.01
			.593	.029	-.285	.014	.111	.027	

Table 7: MSE for Best Performing Models

Method	Moments	T	N	a	b	δ	p
GT	M4	100	500	.1444	.3364	.1369	.3225
IND	M2	100	500	.5300	.1936	.1296	.3226
SMM	m2	100	500	.3136	.1369	.2401	.4272
GT	M4	100	1000	.1444	.3364	.1156	.3136
IND	M2	100	1000	.1764	.2025	.1225	.2993
SMM	m2	100	1000	.2704	.1089	.1849	.3588
GT	M4	100	2500	.1296	.3364	.1089	.2914
IND	M2	100	2500	.1681	.1936	.1225	.2934
SMM	m2	100	2500	.2401	.1225	.1936	.3454
GT	M4	100	INV	.1024	.2401	.1024	.2335
SMM	m2	100	INV	.1936	.1089	.2116	.2850
PMLE	–	100	–	.0994	.1702	.0865	.1837
GT	M4	200	500	.0625	.1444	.0625	.1415
IND	M2	200	500	.1156	.1444	.0841	.2065
SMM	m2	200	500	.1296	.0324	.0961	.1872
GT	M4	200	1000	.0529	.1225	.0576	.1260
IND	M2	200	1000	.1089	.1024	.0729	.1834
SMM	m2	200	1000	.1225	.0289	.0900	.1726
GT	M4	200	2500	.0529	.1089	.0529	.1176
IND	M2	200	2500	.0961	.0900	.0625	.1618
SMM	m2	200	2500	.1089	.0256	.0784	.1519
GT	M4	200	INV	.0529	.1089	.0484	.1165
SMM	m2	200	INV	.0841	.0400	.0900	.1286
PMLE	–	200	–	.0457	.0741	.0457	.0952

MSE for a , b , and δ is the mean squared error computed as the average of the squared difference between the estimate and its true value (times 100). The MSE for p is the average of the squared difference (times 100) between prices based on the true parameters and the estimated parameters, with errors drawn from the same seed.

Figure 1
Density from 500 Monte Carlos and 100 obs

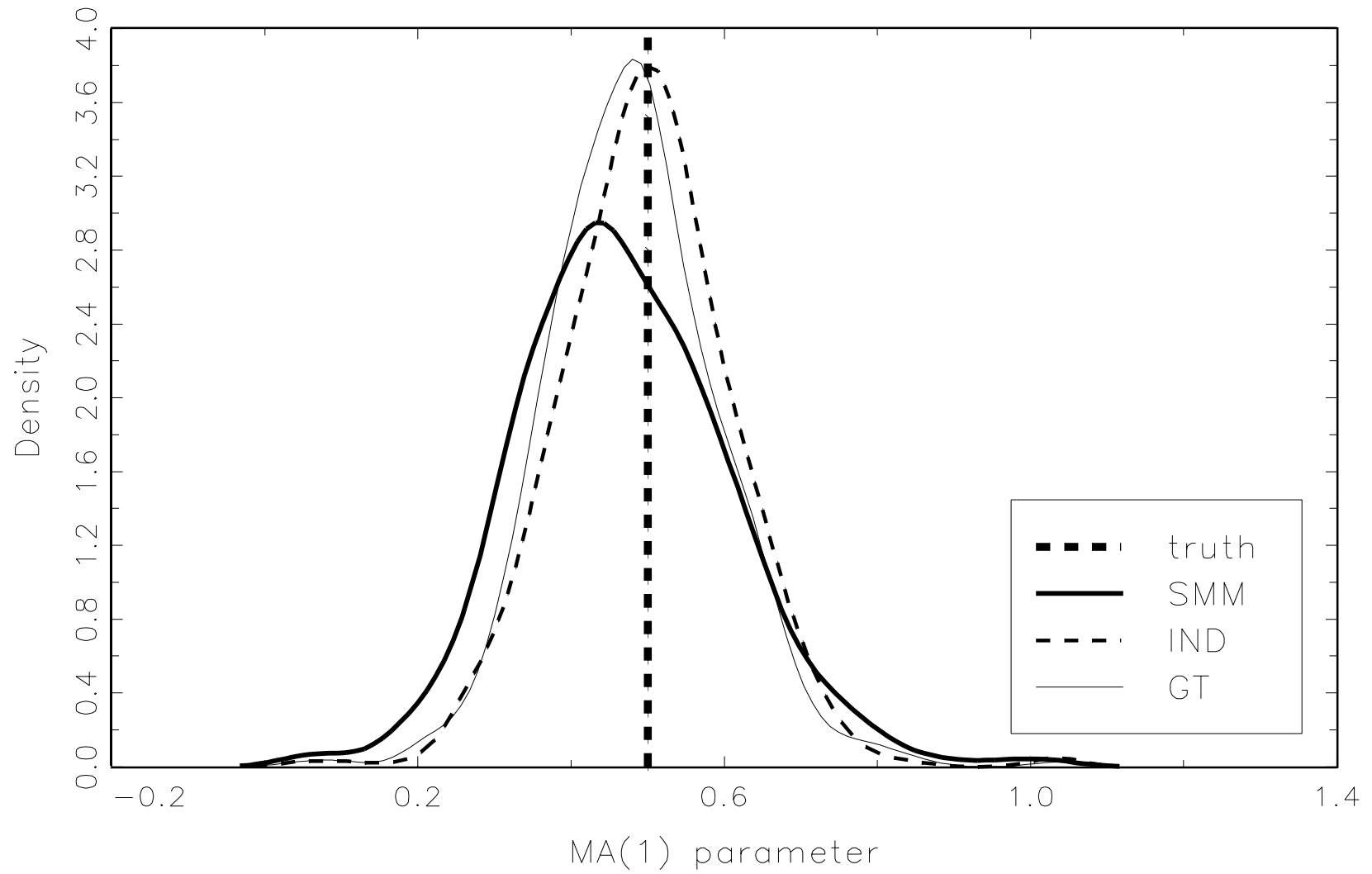


Figure 2
Density from 500 Monte Carlos and 200 obs

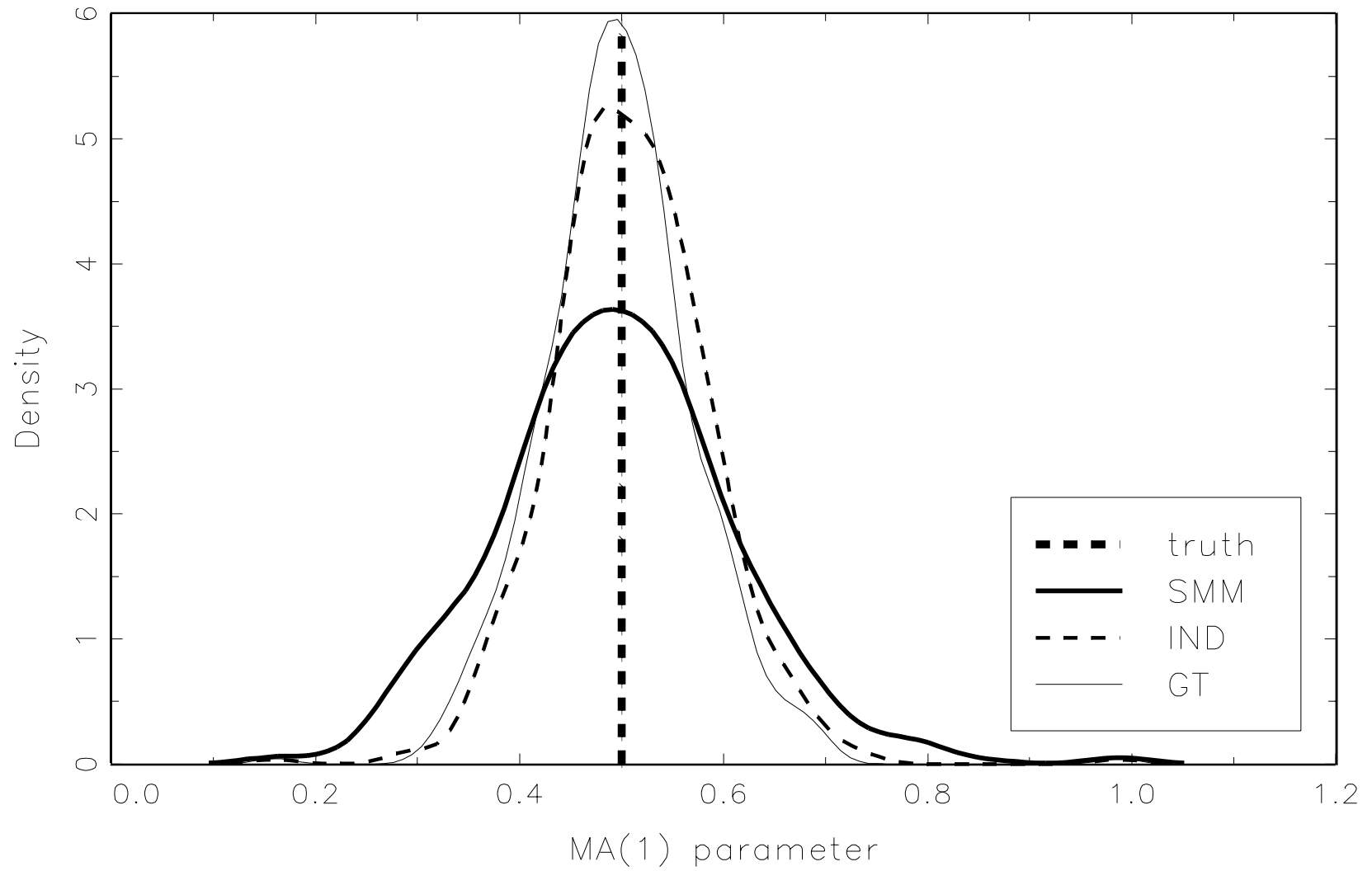


Figure 3: Densities of PMLE estimates from 500 Monte Carlo and 100 obs

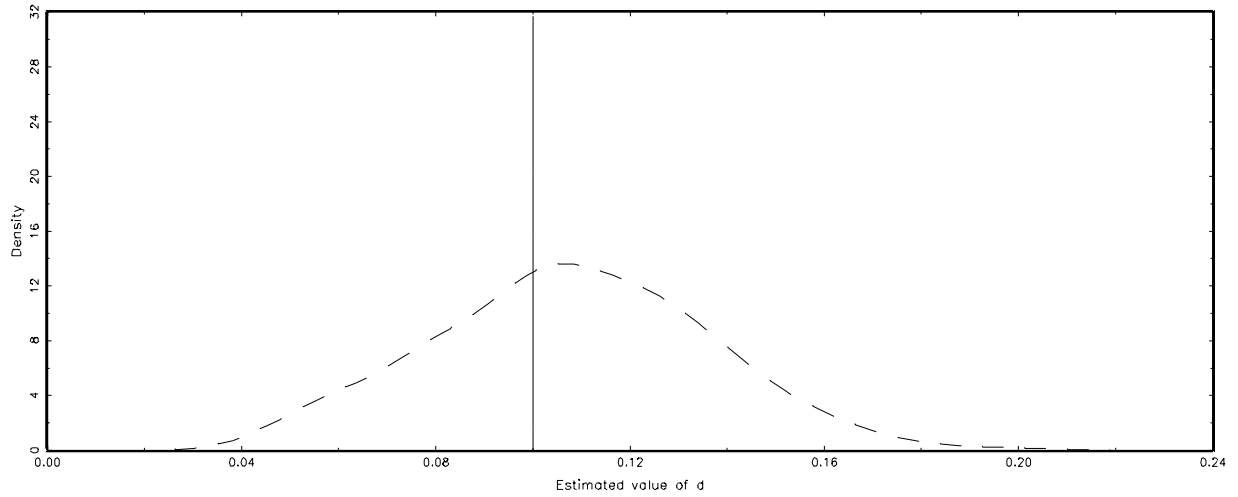
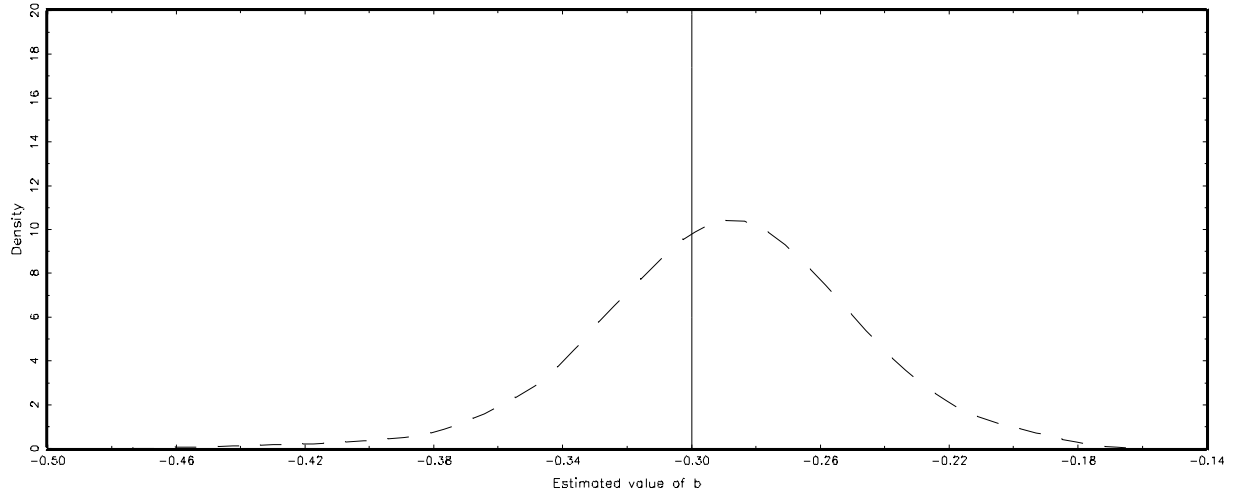
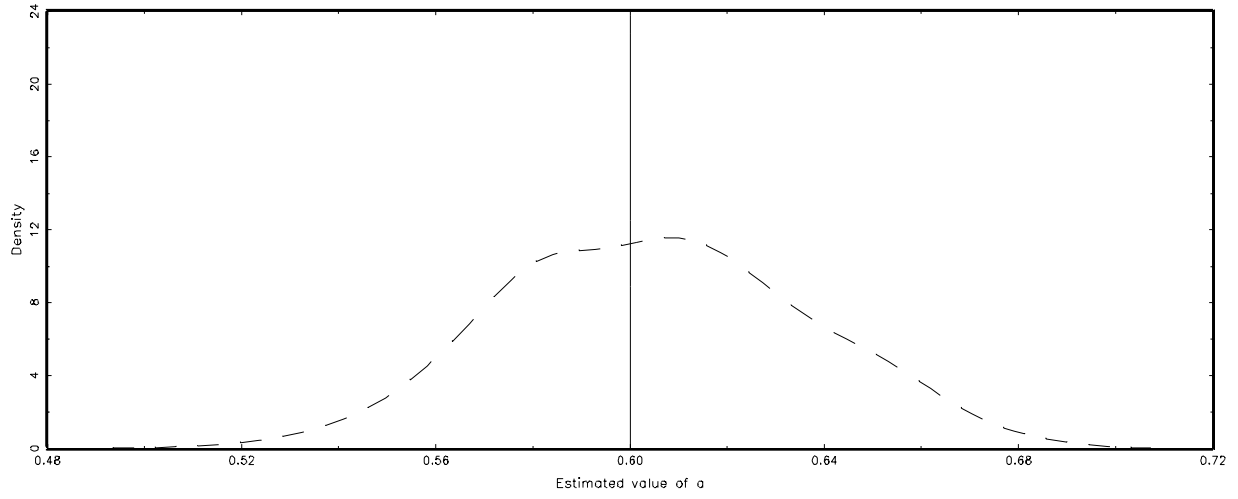


Figure 4: Densities of PMLE estimates from 500 Monte Carlo and 200 obs

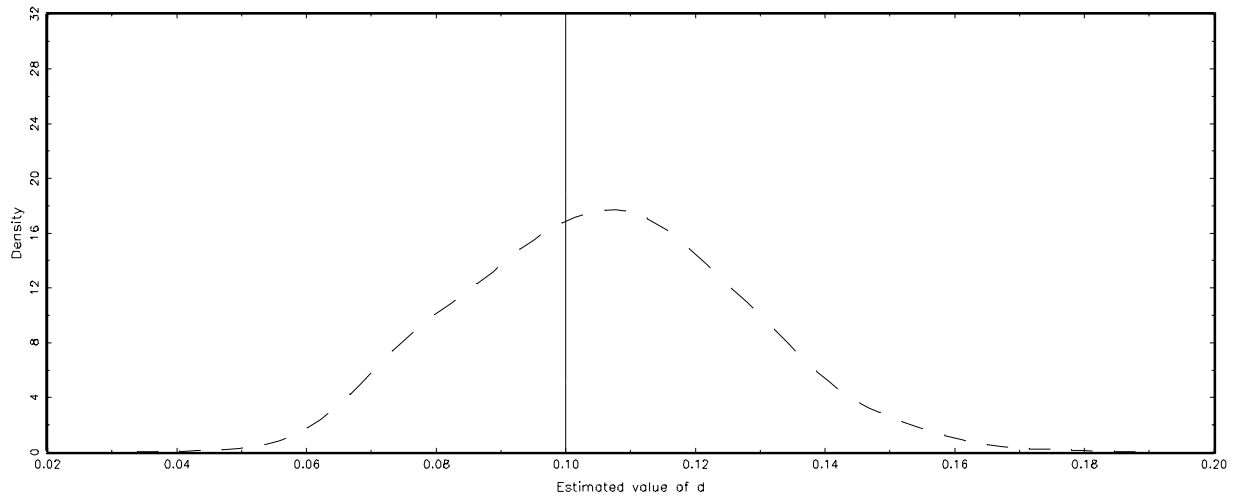
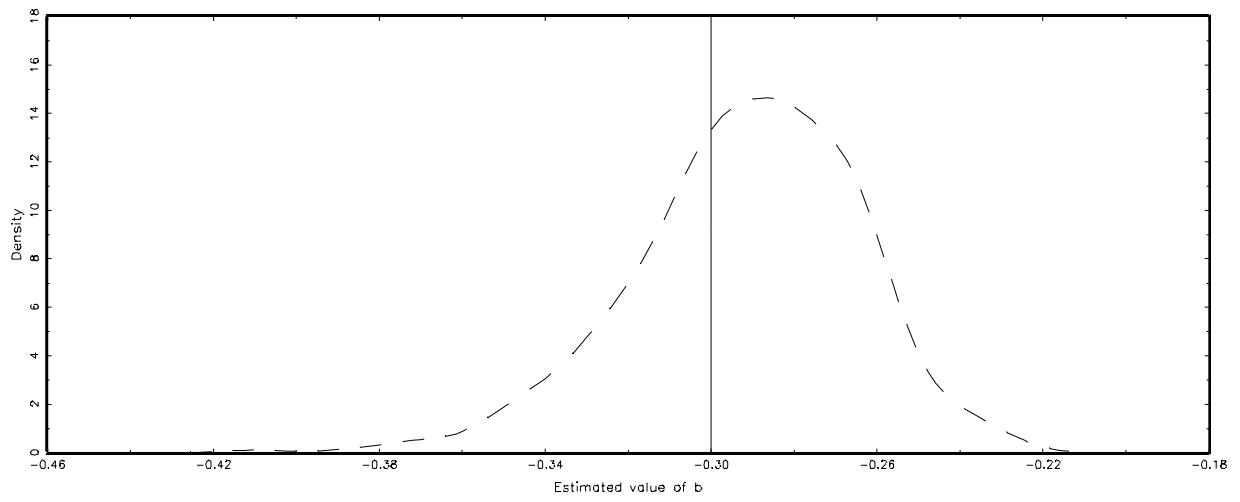
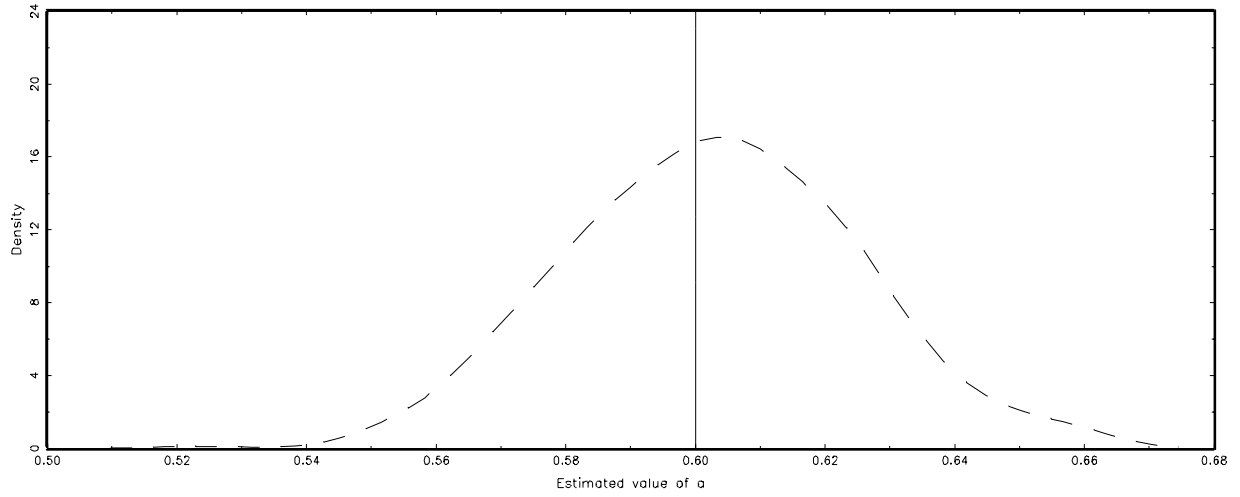


Figure 5
Density from 500 Monte Carlo and 100 obs

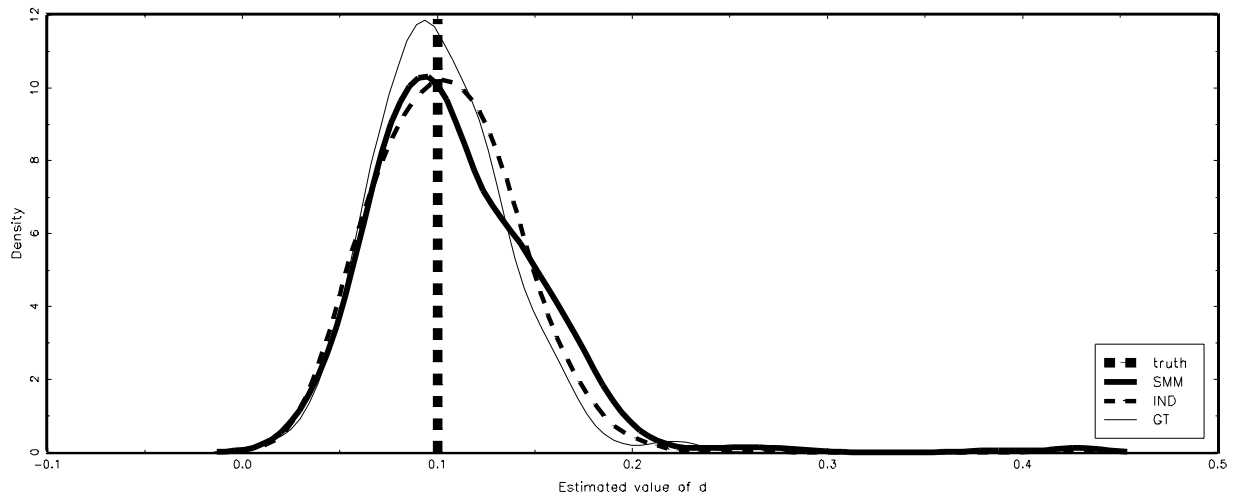
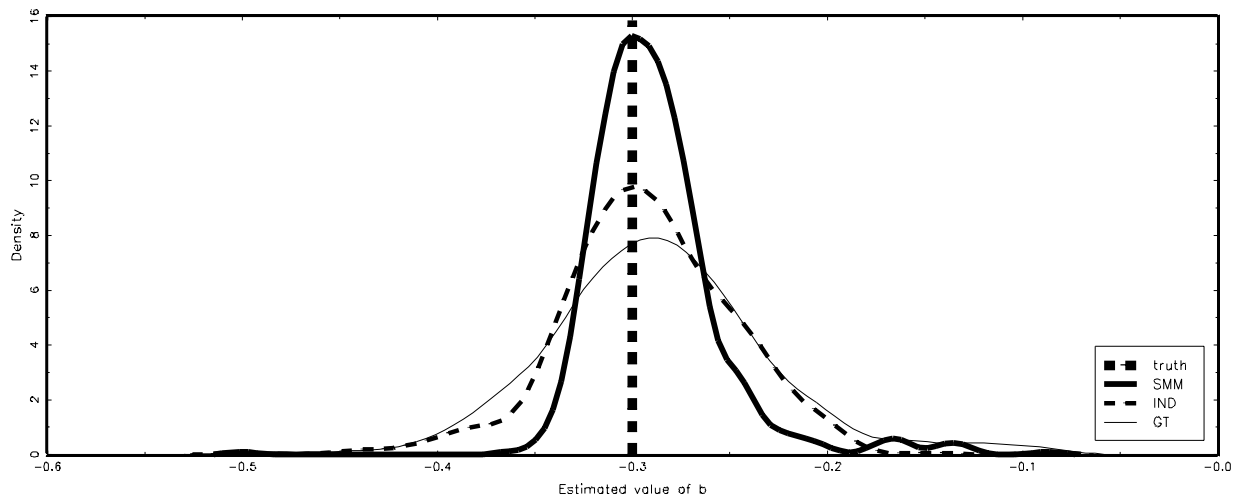
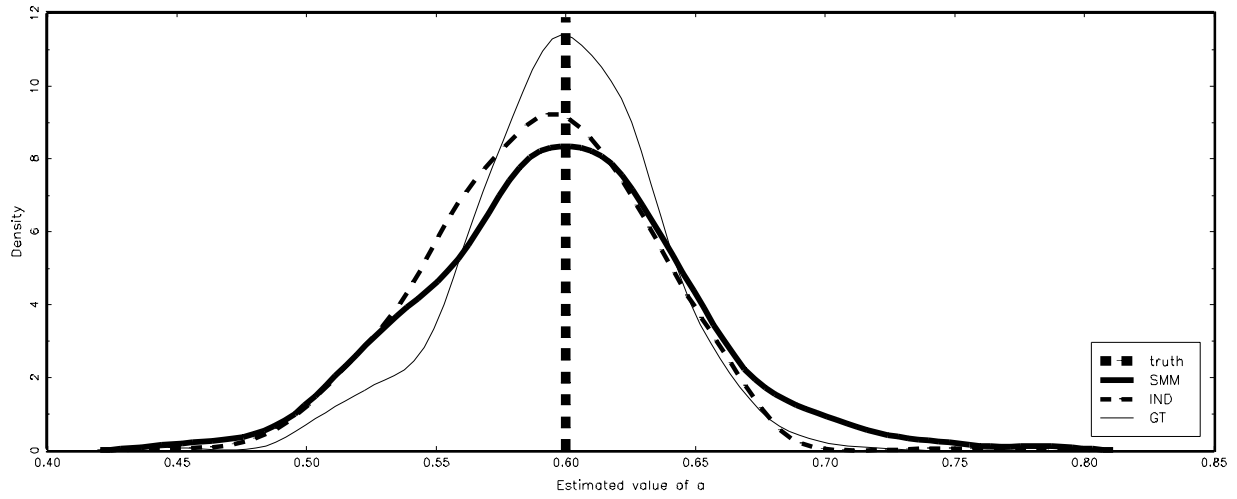


Figure 6
Density from 500 Monte Carlo and 200 obs

