JOURNAL OF APPLIED ECONOMETRICS, VOL. 10, S9–S40 (1995)

ESTIMATING A NONLINEAR RATIONAL EXPECTATIONS COMMODITY PRICE MODEL WITH UNOBSERVABLE STATE VARIABLES

ANGUS DEATON
Woodrow Wilson School, Princeton University, Princeton, NJ 08544, USA

GUY LAROQUE
INSEE, 13 Boulevard Gabriel Péri, 92241 Malakoff Cedex, France

SUMMARY

This paper is concerned with the estimation of a model in which a possibly serially correlated stochastic process, the 'harvest' of an agricultural commodity, generates a competitive price in a market comprising both final consumers and risk-neutral speculators who can store the commodity at a cost in the anticipation of profit. Because storage cannot be negative, the relationship between prices and harvests is inherently nonlinear and is an unpromising candidate for a linear-quadratic model, or for linearization more generally. Instead, we calculate numerically a policy function in which price is a function of two unobservable state variables, the harvest and current availability, and we use the result to fit the price data.

1. INTRODUCTION

In recent years there has been increasing interest in economic models that rest on the computation of policy functions as solutions to dynamic programming problems and the analysis of their associated dynamic equilibria. The real business cycle models of macroeconomics and finance are the most notable examples, but models with similar conceptual structures are also found in the industrial organization literature (see, for example, Pakes, 1994, and Rust, 1992, 1994). Much of the real business cycle literature contents itself with model calibration and with simulation of dynamics, and parameter estimation when it is addressed at all is usually undertaken using limited information techniques such as generalized method of moments (GMM) applied to Euler equations. There exist a number of applications, most notably Rust (1987), where the solutions of dynamic *discrete* choice programs are fitted to the data, but there is little experience to date with estimating the solution of dynamic programs where the choice variable is continuous. The practical calculation of explicit numerical solutions of dynamic programs is particularly burdensome when behaviour depends on several state variables so that the policy functions have multiple arguments. In such cases, including the one considered in this paper, there may be more state variables than variables that are observed, a situation that presents considerable practical difficulties for estimation.

This paper is concerned with a model in which a possibly serially correlated stochastic process, the 'harvest' of an agricultural commodity, generates a competitive price that clears a market comprising both final consumers and risk-neutral speculators who can store the commodity at a cost in the anticipation of profit. Because storage cannot be negative—

CCC 0883–7252/95/0S0S09–32

Received October 1994
Revised June 1995

commodities cannot be consumed before they exist—the relationship between prices and harvests is inherently nonlinear and is an unpromising candidate for a linear-quadratic model, or for linearization more generally. Price is a continuous variable, and so the problem is not one that can be handled by discrete choice techniques. Analysis of Euler equations is useful (see Deaton and Laroque, 1992) but works only when harvests are i.i.d. not serially correlated, and even then leaves key parameters unidentified, without which it is impossible to test whether the model replicates the main features of the data. The difference between i.i.d. and serially correlated harvests is important in another way. In the i.i.d. case, price, which is the only variable that we observe, depends on a single state variable, current availability, defined as the sum of the current harvest and storage from previous periods. The relationship between price and availability is the policy function that must be computed, and once this is done, we have a direct link between price and the unobservable availability. Once the harvest is serially dependent, the price is a function of two state variables, availability and the current harvest. The latter contains information on future harvests and must therefore guide the current storage decision. Policy functions with two state variables are more difficult to compute than policy functions with only one, and the fact that neither state variable is observable clouds the link between the data and the theoretical model.

The model discussed in this paper can be modified or interpreted in a number of ways that would change both its economics and its econometric implementation. For example, the driving shock, which we refer to as the 'harvest', can be interpreted as the difference between a supply and a demand shock, so that we are not assuming that demand is deterministic. It is also possible to allow for deterministic or stochastic trends in both supply and demand, provided their difference is stationary. That said, one of the main purposes of this work is to investigate whether a simple model of linear demand supplemented by speculative storage can deliver the highly autocorrelated commodity prices in a supply-dominated model where harvests are either i.i.d. or modestly autocorrelated. Our assumption that there are no data on availability is also a modelling decision, albeit one that is standard in much of the commodity price literature where the aim has been to use the storage model to deliver predictions about the time-series behaviour of prices. Information on production, consumption, and inventories is usually of relatively low quality, though if adequate data were available, the econometric modelling would be quite different from that presented here, and in some ways more straightforward.

Another important feature of our work is that storage is speculative, and we do not admit the possibility that consumers or manufacturers hold commodity stocks for purposes of convenience. The importance of a convenience yield from commodity storage has a long and disputed history in the commodity price literature (see, for example, Williams and Wright, 1991). On the one hand, convenience yield can, in principle, explain why stocks are held in the expectation of loss, why stockouts are rare, and why prices are strongly autocorrelated. In the extreme case, where there are infinite costs of running out of stock, there will never be stockouts, and the econometric modelling is much simplified over that discussed in this paper (see Miranda, 1994; Rui and Miranda, 1995). However, the precise mechanism through which the presence of commodity stocks affects production has never been clearly spelled out—for example, through an inventory management or Ss-model—and it remains unclear whether the magnitude of convenience yield required to explain the data is plausible in reality. Our own approach, as in Williams and Wright, is to discover whether or not speculative storage by itself is capable of explaining the time series properties of commodity prices. If not, convenience yield is among the alternative explanations that should be seriously explored.

This paper has two immediate purposes. The first is to provide a more complete technical discussion of the methodology used to obtain the results in Deaton and Laroque (1996), where

the main focus is on our substantive empirical findings for the commodity price model. The second is to provide a general discussion of the computational and econometric issues that arise for this problem since many of the issues are not specific to this model but would arise in most related exercises in confronting data with the solutions to dynamic programs. The paper is laid out as follows. Section 2 provides a brief but self-contained exposition of the commodity price model, focusing on the specifications that are used in the estimation; more detailed treatments of the model are contained in Deaton and Laroque (1992, 1996) and in Chambers and Bailey (1992). This section also explains the general approach to estimation. Section 3 discusses the 'inner loop' of the calculations, where the policy functions are numerically approximated. As outlined in the previous paragraph, estimation is a good deal more straightforward when the harvests are assumed to be i.i.d., and estimation in this case is discussed in Section 4. Section 5 turns to the most general case where the harvests follow a first-order autoregression, and where the estimation becomes more difficult and more experimental.

2. THE COMMODITY PRICE MODEL, THE DATA, AND GENERAL ECONOMETRIC METHOD

2.1. The Model

Consider the following simplified model of an agricultural commodity. In each year, there is an exogenous harvest z_t which follows a stochastic process characterized (in the most general case that we consider) by a transition function $\Lambda(z, Z)$, which gives the probability that the harvest in year $t+1$ is less than or equal to Z when the harvest in period t is z. The harvest shocks come from outside the model, are determined by the conditions of agricultural production, and are what ultimately drives the behaviour of prices. We shall work with two cases, one where z is i.i.d. and one where z follows a first-order linear autoregressive process; although the former is a special case of the latter, it will be convenient to treat them separately.

There are two kinds of actors in the market: final consumers and risk-neutral speculators. The consumers have inverse demand functions that we take to be linear, so that if there were no storage and the harvest z were consumed in each period, price would be given by

$$P(z) = a + bz \tag{1}$$

where a and $b<0$ are parameters. Given equation (1), and if consumers were the only demanders in the market, we could directly infer the behaviour of prices from the behaviour of the harvests. In particular, prices will be i.i.d. if harvests are i.i.d., and prices will follow an AR(1) if the harvest follows an AR(1). A good reason for choosing the linear demand (1) is the reduction in the dimensionality of the parameter space that is provided by the identification proposition discussed below. While such identification results cannot be provided in general, there are other (nonlinear) demands—such as negative exponential—for which similar results hold.

Inventory holders demand the commodity in order to transfer it into the next period, and will do so whenever they expect to make a profit over the storage and interest costs. The technology of storage takes the simple form of proportional deterioration during storage, so that if I units are stored, $(1-\delta)I$ are available at the beginning of the next period; speculators also have to pay a real interest rate r per period on the value of commodities stored. Define x, the 'availability' or 'amount on hand' as the sum of the harvest and inherited inventories, so that

$$x_t = (1-\delta)I_{t-1} + z_t \tag{2}$$

The crucial nonlinearity in this specification is introduced by the requirement that inventories be non-negative. Permitting the market to borrow the commodity from the future would be to permit the consumption of bread from grain that has not yet been grown. Given this, together with the assumptions that speculators are risk neutral and have rational expectations, current and expected future prices must satisfy

$$p_t = \max\left[P(x), \frac{1-\delta}{1+r} E_t p_{t+1}\right] \tag{3}$$

The second term in brackets is the expected value of one unit stored, after allowing for storage costs, conditioned on the speculator's current information, assumed to be the current harvest and the current availability. The first term is the price that would hold if current availability were sold to consumers, and this will be the actual price if it is greater than the net expected future price. This is the situation in which speculators would hold negative inventories were it possible to do so. If the net expected price is greater than the price when no inventories are held, speculators exploit the arbitrage opportunities and drive up the current price until current price and their expectation of the net future price are equal.

In this market-based model, equation (3) comes from an arbitrage condition, but, as has long been recognized in the commodity price literature, it can also be regarded as an optimality condition for storage in a social planning problem (Samuelson, 1970). It is also formally identical to the optimality condition on intertemporal consumption choice for a consumer who can save but who has no access to borrowing (see Deaton, 1991). More generally, condition (3) would correspond to the Euler equation in real business cycle or other dynamic models, and in some simple cases can serve as the basis for estimation, an issue to which we shall return below.

Provided certain conditions are met, most notably that $r+\delta>0$, there exists a solution to equation (3) in terms of a price function

$$p = f(x, z) \tag{4}$$

where $f(x, z)$ is the unique monotone decreasing in its first argument solution to the functional equation

$$f(x,z) = \max\left(\frac{1-\delta}{1+r}\int f(z' + (1-\delta)(x - P^{-1}(f(x,z))), z')\Lambda(z, \mathrm{d}z'), P(x)\right) \tag{5}$$

(see Deaton and Laroque, 1992; Chambers and Bailey, 1996). Except in very special cases, this equation does not permit an analytical solution, and must be solved numerically, techniques for which are the main topic of the next section.

It is useful to restate equation (5) for the two cases with which we shall be concerned. When the harvests are i.i.d., the transition probabilities are independent of z, so that $f(x, z)$ is also independent of z. The price function is then characterized by the simpler form of equation (5):

$$f(x) = \max\left(\frac{1-\delta}{1+r}\int f(z' + (1-\delta)(x - P^{-1}(f(x)))) \,\mathrm{d}\Lambda(z'), P(x)\right) \tag{6}$$

where, in an abuse of notation, $\Lambda(.)$ is the univariate c.d.f. of the harvest process. The second case is when the harvest process follows a linear autoregression

$$z_{t+1} = \rho z_t + \varepsilon_{t+1} \tag{7}$$

for $-1 < \rho < 1$, and the ε's are i.i.d. with mean μ, say. Specializing from equations (5) to (7) gives

$$f(x, z) = \max[G(x, z), P(x)] \tag{8}$$

where $G(x, z)$ is the expected price net of storage costs,

$$G(x, z) = \frac{1-\delta}{1+r} \int f(\varepsilon + \rho z + (1-\delta)(x - P^{-1}(f(x, z))), \varepsilon + \rho z)\, \mathrm{d}\Phi(\varepsilon) \tag{9}$$

and Φ is the univariate distribution of ε.

Once the functions have been computed, they can be used to characterize the behaviour of the stochastic processes. In particular, note that, by equation (2), availability evolves according to

$$x_{t+1} = (1-\delta)(x_t - P^{-1}(f(x_t, z_t))) + z_{t+1} \tag{10}$$

and the harvest z_t evolves according to the transition function $\Lambda(z, Z)$, which in our case means that it is either i.i.d. or follows the autoregression (7). In either case, we have a bivariate stochastic process in the vector (x_t, z_t), both elements of which are unobserved, and with which there is the associated observable price given by equation (4). In the i.i.d. case, the harvest does not appear in the price function conditional on availability, so that availability is the only state variable, to which price is monotonically related.

2.2. The Data and General Econometric Method

Our data are annual time series of commodity prices. Although there exists some information on harvests and availability for some commodities for some time periods, it is not possible to construct consistent quantity series for the long run of years for which prices are available. With only price data, we are limited in the number of parameters that we can hope to recover from the data. In the formulation described above, apart from the distribution of the harvest innovations there are four free parameters in the i.i.d. case (the depreciation rate δ, the real interest rate r, and the slope and intercept of the inverse demand function a and b), plus a fifth in the serially correlated case, the autoregressive coefficient ρ. It turns out that, when there are no data on the harvests or on the stocks and when the demand function is linear, it is impossible to recover separately the demand function and the mean and standard errors of the disturbances. The following proposition is proved in the Appendix of Deaton and Laroque (1996):

Proposition (identification when only prices are observed):

Consider an economy E, with real interest rate r and inventory depreciation parameter δ, where the harvest shocks z have zero mean, unit variance and autocorrelation ρ, and where the inverse demand function $P(x)$ is linear:

$$P(x) = a + bx$$

Any other economy, with the same r and δ, with harvest shocks $\tilde{z} = \mu + \sigma z$ and inverse demand function

$$\tilde{P}(x) = \left(a - \frac{b\mu}{\sigma}\right) + \frac{b}{\sigma} x$$

has the same rational expectations price process as the base economy E.

In view of this result, we normalize the innovation of the harvest process to be of mean zero and of variance one; the standard deviation of the harvest z is thus $1/\sqrt{1-\rho^2}$. For simplicity, we have taken the distribution to be normal—although see Section 3.1—and also fixed the real interest rate at 5%.

Our general approach is to use the theory to construct one-period-ahead conditional expectations and variances of prices, and to match these theoretical conditional moments to the actual data. There are two elements in this process: the construction of the moments and the matching to the data. The latter is done using pseudo maximum likelihood estimation (PMLE) (see Gourieroux *et al.*, 1984). Write the one-period-ahead conditional mean and variance of price as $m(p_t)$ and $s(p_t)$, so that

$$\begin{aligned} m(p_t) &= E(p_{t+1} \mid p_t) \\ s(p_t) &= V(p_{t+1} \mid p_t) \end{aligned} \tag{11}$$

and define (twice the log of) the pseudo-likelihood function by

$$2 \ln L = \sum_{1}^{T-1} \ln l_t = -(T-1)\ln(2\pi) - \sum_{1}^{T-1} \ln s(p_t) - \sum_{1}^{T-1} \frac{(p_{t+1} - m(p_t))^2}{s(p_t)} \tag{12}$$

This function is maximized with respect to the parameters of the model. Equation (12) would yield the exact likelihood function if prices were normally and heteroscedastically distributed conditional on lagged price. When there is storage, we would not expect prices to be normally distributed, even if the harvests were normal, but the parameter estimates will nevertheless be consistent, in spite of the fact that both the conditional expectation and conditional variance are non-differentiable in p_t so that the likelihood may be non-differentiable in the structural parameters (see Laroque and Salanie, 1994).

The variance–covariance matrix is calculated in the standard way: for the k-vector of parameters β, define the $(T-1)$ by k matrix G and the k by k matrix J by

$$g_{ti} = \frac{\partial \ln l_t}{\partial \beta_i}; \qquad J_{ij} = \frac{\partial^2 E \ln L}{\partial \beta_i \partial \beta_j} \tag{13}$$

Then the asymptotic variance covariance matrix of the parameters is given by

$$V = J^{-1}(G'G)J^{-1} \tag{14}$$

The construction of the moments (11) will be described in some detail in Sections 4 and 5 below. The calculations are somewhat different depending on whether or not the harvests are i.i.d., and, as usual, the i.i.d. case is a good deal simpler. Suppose that in this case we have calculated the (univariate) price function $f(x)$. Since it is monotone it can be inverted, so that with each price p_t we can associate the corresponding availability $x_t = f^{-1}(p_t)$. From the evolution of availability in the i.i.d. case, we have from equation (10)

$$x_{t+1} = (1-\delta)(f^{-1}(p_t) - P^{-1}(p_t)) + z_{t+1} \tag{15}$$

so that the next period's availability is the next period's harvest plus inventories from this period, which are the (depreciated) difference between current availability and current consumption. The next period's price is therefore

$$p_{t+1} = f(x_{t+1}) = f[(1-\delta)(f^{-1}(p_t) - P^{-1}(p_t)) + z_{t+1}] \tag{16}$$

which, conditional on p_t, contains the single i.i.d. random variable z_{t+1}, so that the conditional moments can be obtained by integration. It is also worth noting that equation (16) provides a direct characterization of the price process as a nonlinear autoregression. In principle, this equation could be used to estimate the model by maximum likelihood, using nonlinear Kalman filtering or some similar technique. One potential problem is that the function $f(.)$, although continuous and monotone, is not differentiable, so that the resulting discontinuity in the Jacobian would require the maximization of a non-continuous likelihood function. (Note that the full-information maximum likelihood estimation of the simulated commodity price model in Miranda, 1994, is in the context of a model without stockouts whose policy function and likelihood function are everywhere differentiable.)

When the harvests are i.i.d., it is also possible directly to characterize the one-period-ahead conditional expectation of price in equation (11). In the i.i.d. case, price is a function of only availability, $f(x)$, and this function has the property that, for x less than some critical value x^*, with associated critical price $p^* = f(x^*)$, $f(x) = P(x)$, while for $x > x^*$, $f(x) > P(x)$ (Deaton and Laroque, 1992, Theorem 1.) When availability is lower than the critical level, nothing is stored, and prices are what they would be in the absence of storage. When this happens, the next period's price p_{t+1} is simply $f(z_{t+1})$ (see equation (16)) so that expected price is the expectation of $f(z_{t+1})$, which, since the harvest is i.i.d., is a constant. Summarizing,

$$E(p_{t+1}|p_t) = \int f(z)\, \mathrm{d}\Lambda(z), \quad \text{for } p_t \geq p^* \tag{17}$$

When prices are below the critical level p^*, so that availability is above x^*, there will be storage, so that the net expected future price must equal the current price. Hence

$$E(p_{t+1}|p_t) = \frac{1+r}{1-\delta} p_t \quad \text{for } p_t \leq p^* \tag{18}$$

Combining equations (17) and (18) gives a simple piecewise linear form for the conditional expectation:

$$E(p_{t+1}|p_t) = \frac{1+r}{1-\delta} \min[p_t, p^*] \tag{19}$$

As far as we are aware, no similarly convenient form exists for the conditional variance in the i.i.d. case, nor, as we shall see below, for either conditional mean or conditional variance when the harvests are serially dependent.

In this i.i.d. case, equation (19) can be used to estimate the model using GMM techniques; the difference between p_{t+1} and its prediction is orthogonal to information dated t and earlier, so that current and lagged prices make natural instruments for the GMM estimation. This estimation is carried out by Deaton and Laroque (1992) on annual data and by Chambers and Bailey (1996) on monthly data. Note however that equation (19) does not identify r separately from δ, nor does it make any reference to the characteristics of the driving process, except through the choice of instruments. This is a disadvantage if we want to identify parameters of the process, but an advantage if we know nothing about it, and want to make inferences that are robust to our ignorance, the usual trade-off in using a limited information over a full-information method. Equation (19) also does not reflect the commodity price model's theoretical predictions about conditional variances; for example, that when the inverse demand function is convex, conditional variance is monotone non-decreasing in prices. The PML estimation outlined above does use this information, but in contrast to full-information

maximum likelihood, makes no use of the theoretical predictions about conditional skewness and kurtosis.

More generally, note that Euler equation estimation does not seem to be readily applicable to the commodity price problem *except* in the i.i.d. case. Although equation (3) can be interpreted as an Euler equation for optimal allocation in a social planner's problem, it is not directly amenable to GMM estimation because of the nonlinearity of the *max* function. Except where there is some special feature of the problem that leads to expressions like (19), the presence of the expectation inside the nonlinear function precludes the construction of the standard GMM criterion function that is used to derive consistent estimates.

In the serially dependent case, it is not possible to make the calculations (15) to (19) because the price is now a function of the two variables z and x. In consequence, we cannot associate a unique availability with each price in the data, and it is no longer the case that prices can be characterized by any first-order process like equation (16). Nor does the conditional autoregression of prices satisfy the piecewise linear form (19) when the harvests are serially dependent. As in the i.i.d. case, there is a critical price above which inventories are zero and prices are given by $P(x)$, but this critical price is not constant as it is in the i.i.d. case but is a function of the harvest z. As a result, equation (19) holds only if we condition on *both* lagged price and the harvest,

$$E(p_{t+1}|p_t, z_t) = \frac{1+r}{1-\delta} \min[p_t, p^*(z_t)] \tag{20}$$

Since we do not observe z_t, this expression is a good deal less useful than is equation (19). It does not permit GMM estimation, and it cannot be used to calculate $E(p_{t+1}|p_t)$ without prior calculation of both the functions $p^*(z)$ and the distribution of z_t conditional on p_t. In order to calculate the conditional means and variances (11), we must therefore follow a different strategy in the serially dependent from the i.i.d. case. This strategy is discussed in detail in Section 5; in outline we proceed as follows.

For each price in the data p_t we calculate a set of values of (x_t, z_t) that are consistent with that price. For each element of the set, we use the transition function to calculate the transition probabilities (densities) to each possible (x_{t+1}, z_{t+1}) in period $t+1$, and thus to the corresponding price $p_{t+1} = f(x_{t+}, z_{t+1})$. In order to put relative weights on these various p_{t+1}, we assume that the data in period t are drawn from the unique invariant limit distribution of the bivariate process on harvests and availability. We compute this distribution and derive from it the conditional probabilities associated with each possible (x_t, z_t) consistent with p_t. Weighting the associated values of p_{t+1} yields the conditional expectation and variance of p_{t+1}, for insertion in the PML criterion. This is certainly not the only possible method of estimation. For example, we might compute the moments of prices conditional on the whole history using a filtering procedure. We might also use the moments differently, for example in a GMM scheme, or more radically, we could apply the moment matching technology of Gallant and Tauchen (1994). These are alternatives that we plan to investigate in further research.

3. COMPUTING THE PRICE FUNCTIONS

The core of the calculations is the solution of the functional equations to derive approximations for the functions relating price to harvest and availability. Although the price functions are different in the i.i.d. and dependent cases, the calculations have many common features and

are dealt with together in this section. This material is also relatively familiar from similar calculations in the literature.

The need to solve for a policy function—here the price function—is what distinguishes this sort of problem from a standard one in nonlinear estimation. As in the standard problem, we set up a criterion (pseudo-likelihood) function in terms of the parameters and choose the parameters to maximize it. In the current calculation, the evaluation of the criterion function cannot be carried out until we have calculated the unknown function $f(x, z)$, which itself involves a nonlinear computational problem. In this paper, the approximation is done by calculating the value of $f(x, z)$ at a finite number of 'grid' points and interpolating between them. Essentially, we construct a 'table' for the function, and use the tabulated values as benchmarks for calculating other values. Once the grid points have been chosen, we start with some guess for the function, and then use the functional equation (5) or its specializations (6) or (8) and (9) to improve our guess in a systematic way. Before we discuss these procedures in detail, it is convenient to begin with the specification of the harvest process.

3.1. The Discretization of the Harvest Process

Our starting point for the process z_t is the standard normal distribution; recall that in view of the Proposition, we are free to standardize on zero mean and unit variance, but the choice of normality is a substantive one. To calculate the price functions, we shall repeatedly have to take expectations of nonlinear functions of z_t which requires computationally expensive numerical integration. The calculations can be speeded up a great deal if the harvest is a discrete rather than continuous process, and this subsection explains how this is done so as to approximate the original standard normal. The procedures are similar to those proposed by Tauchen (1986) and Tauchen and Hussey (1991).

We wish to replace the unit normal harvest with one that is restricted to N discrete points $Z = (Z_1, \ldots, Z_N)$; in the calculations in this paper $N = 10$, but we retain the general notation. Start by defining an $N+1$ vector $\theta = (\theta_1, \theta_2, \ldots, \theta_{N+1})$ with $\theta_1 = -\infty$ and $\theta_{N+1} = \infty$, and satisfying

$$\Phi(\theta_{j+1}) - \Phi(\theta_j) = 1/N, \qquad j = 1, \ldots, N+1 \tag{21}$$

where $\Phi(.)$ is the cdf of the unit normal. The vector θ divides up the support of the normal distribution into N equiprobable intervals. The N discrete harvest values are then chosen to be the conditional means within each interval, i.e.

$$Z_i = \frac{\int_{\theta_i}^{\theta_{i+1}} z \, d\Phi(z)}{\Phi(\theta_{i+1}) - \Phi(\theta_i)} = \frac{\phi(\theta_i) - \phi(\theta_{i+1})}{\Phi(\theta_{i+1}) - \Phi(\theta_i)} \quad i = 1, \ldots, N \tag{22}$$

and ϕ is the pdf of the unit normal. For the i.i.d. case, we are done; the harvest in each period is a random draw from the N values Z_i, each selected with equal probability $1/N$. When N is 10, as here, the Z-values are (±1·755, ±1·045, ±0·677, ±0·386, ±0·126). With probabilities 0·10, these values have mean zero but standard deviation 0·964, not 1 as in the underlying normal; such approximations cannot reproduce *all* features of the underlying distribution, an issue to which we return below.

When the harvests are serially dependent, we use exactly the same Z-values defined in equation (22), but the probabilities of selection are affected by the harvest in the previous period and the transition probabilities are chosen so as to mimic the properties of an AR(1) with normal innovations. In a continuous normal autoregression, the joint distribution of the harvests

in any two adjacent periods would be bivariate normal with zero means, unit variances, and covariance equal to the autocorrelation parameter and we use this distribution to calibrate the transition probabilities between the Z-values. In particular, for the reference continuous distribution with autocorrelation parameter ρ, the probability that the harvest in $t+1$ lies in the interval (θ_{i+1}, θ_i) conditional on the harvest in t lying in (θ_{j+1}, θ_j) is given by

$$T_{ij}(\rho) = \frac{\int_{\theta_i}^{\theta_{i+1}} \int_{\theta_j}^{\theta_{j+1}} \phi_B(x_1, x_2; \rho)\, dx_2\, dx_1}{\Phi(\theta_{j+1}) - \Phi(\theta_j)} \tag{23}$$

where $\phi_B(x_1, x_2; \rho)$ is the standard bivariate normal density function

$$\phi_B(x_1, x_2; \rho) = [2\pi\sqrt{1-\rho^2}]^{-1} \exp[-(x_1^2 - 2\rho x_1 x_2 + x_2^2)/2(1-\rho^2)] \tag{24}$$

The integration in equation (23) is done using the routines in GAUSS that provide the tail areas of bivariate normal densities. We then use the T-matrix calculated from equation (23) as the transition matrix for the discrete stochastic process on Z, so that our approximation to the AR(1) is defined by the N-vector Z together with the transition probabilities

$$\text{prob}(Z_t = z_i \mid Z_{t-1} = z_j) = T_{ij}(\rho) \tag{25}$$

As in the univariate case, these transition probabilities and associated Z-values do not give a perfect approximation. For an autoregressive parameter of 0·7, Figure 1 shows the conditional expectation $E(z_t \mid z_{t-1} = Z) = T(\rho)'Z$ together with ρZ, which is the conditional expectation for the continuous bivariate distribution; the approximation here is quite close, as it is for other values of ρ. Figure 2 shows the one-period-ahead conditional variances calculated from the discrete approximation, a quantity that is unity in the bivariate autoregression. The illustrations are for the 10-point harvest and with ρ set to values from 0·1 to 0·9. As ρ becomes larger, the discrete conditional variance deviates increasingly from the constant value of unity. Figure 3 shows a 200-period realization of the discrete stochastic process, again with ρ set to 0·7. The fact that we are working with a grid can clearly be seen, particularly when the same value is repeated for more than one period, but otherwise the realizations look very much like a standard linear AR(1).

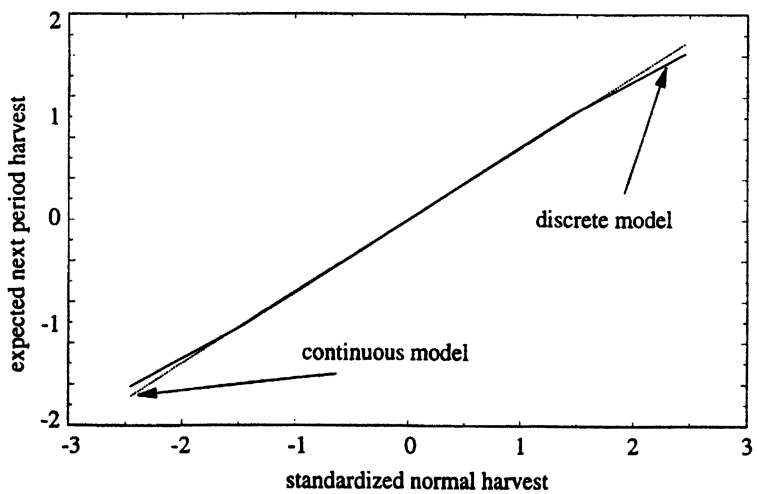

Figure 1. Conditional expectations for continuous and discrete models

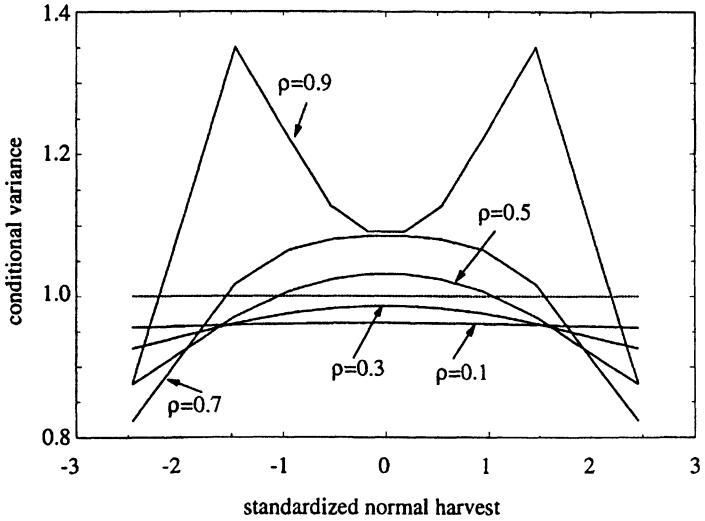

Figure 2. Conditional variance for continuous and discrete models

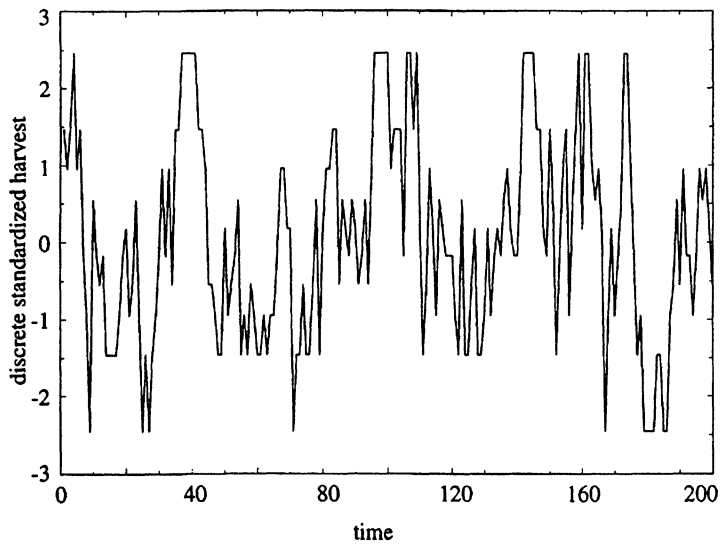

Figure 3. A 200-period simulation of a discrete version of AR(1)

3.2. Computing the Price Function

Given the discrete harvest process and the transition matrix, it is possible to write the price function (5) in a slightly different and more practical form. If we denote by β the ratio $(1-\delta)/(1+r)$, the discreteness of the harvest implies that equation (5) can be written as

$$f(x, Z_i) = \max\left\{\beta \sum_{j=1}^{N} T_{ji} f[Z_j + (1-\delta)(x - P^{-1}[f(x, Z_i)]), Z_j], P(x)\right\} \tag{26}$$

for $i = 1, \ldots, N$. Since the harvest can only take on N distinct values, there are N distinct price functions, each a function of the amount on hand x.

To calculate the price functions in equation (26), we require a numerical approximation which, to some prespecified degree of accuracy and over some range, satisfies the equation. One technique that has been commonly employed in the literature is the use of global polynomial approximations, see for example the symposium edited by Taylor and Uhlig (1990), Judd's (1992) survey, and, in the current context, Williams and Wright (1991, Ch. 2). We can choose a low-degree polynomial approximation for each $f(x, Z_i)$, substitute into both sides of equation (26), and then choose the parameters to minimize the distance between the left- and right-hand sides evaluated at some prespecified grid points. Higher accuracy can be obtained by increasing the number of grid points and the degree of the polynomial. We can also exploit known particular features of the problem, for example that for x less than $x^*(i)$, $f(x, Z_i) = P(x)$, treat the $x^*(i)$ as parameters, and fit the polynomial approximations only for $x > x^*(i)$. Williams and Wright's procedure is of this form. In our (limited) experience, algorithms such as these are fast and accurate, and they are relatively easily computed. However, when we come to estimate the model, we rely on the monotonicity of the calculated price functions in order to obtain inverses (see equations (15) and (16) above) and we found no convenient way of guaranteeing the monotonicity of the polynomial approximations and encountered several failures of monotonicity in practice.

In the face of the difficulty, we abandoned this methodology early in the research in favour of local interpolation along a grid. We work with an equally spaced grid of x-values, $(X_1, X_2, \ldots, X_M)$ which we denote X. Although it would be possible in principle to extrapolate values of $f(x, Z_i)$ beyond the end points in either direction, we regarded such extrapolation as dangerous and took pains to avoid it. While it is desirable to have the smallest possible range of x, so that a given number of grid points gives a finer grid and greater accuracy, the range must be large enough so that, for any relevant parameter values, the calculation never generates values of the amount on hand greater than the maximum value nor less than the minimum value in X. One limit that x cannot exceed is obtained by supposing that the harvest remains indefinitely at its maximum value and that nothing is consumed in which case by equation (10), we would have $x_{t+1} = (1 - \delta)x_t + Z_N$, for maximum harvest Z_N so that x asymptotes to Z_N/δ. At the other extreme, since inventories cannot be negative, the smallest value of x is the minimum harvest Z_1, which is therefore the minimum point on the grid. The use of these values as end points of the grid guarantees that the calculation of the price functions will never generate values outside the grid. (When we come to estimation in the next section, it will be necessary to briefly reconsider the grid points, but the foregoing discussion is sufficient for the calculation of the price functions given any set of parameters.)

Our basic procedure is one of trial and error; a guess is made for each of the N functions, and the guesses improved upon using equation (26). We start from some suitable non-negative, convex, monotone non-increasing function $f_0(x, Z_i)$, and update using equation (26). Suppose, for example, that we choose

$$f_0(x, Z_i) = \max(P(x), 0) \tag{27}$$

and that we update the function using

$$f_{n+1}(x, Z_i) = \max\left\{\beta \sum_{j=1}^{N} T_{ji} f_n[Z_j + (1 - \delta)(x - P^{-1}[f_n(x, Z_i)]), Z_j], P(x)\right\} \tag{28}$$

for $n = 0, 1, 2, \ldots$. Although we can evaluate $f_0(x, Z_i)$ for all values of x using equation (27), we use equation (28) to evaluate $f_1(x, Z_i)$ only at the M points on X. Hence, when we come to

evaluate $f_2(x, Z_i)$, we need some method of interpolation since the right-hand side of equation (28) is not restricted to the grid points. In our initial work (Deaton and Laroque, 1992), we used simple linear interpolation between the grid points, a technique that worked well for simulation. However, the linear interpolation has the effect of introducing non-differentiabilities into the calculated price function at each of the grid points, something that complicates the estimation. We therefore replaced linear interpolation by a standard cubic spline procedure, as, for example, in Press *et al.* (1989). Interpolation by cubic splines is rapid and convenient in practice, especially when the points on the grid are uniformly spaced, and if the interpolations are repeated frequently over the same grid, the most expensive part of the calculation need only be carried out once. The cubic splines also enable us to capture some of the advantages of the polynomial approximations, requiring fewer grid points. Of course, there is once again the risk of non-monotonicity, though this can be avoided by making the grid sufficiently fine, or by using shape preserving spline methods as in Schumaker (1983).

Note that if the innermost $f_n(x, Z_i)$ on the right-hand side of equation (28) were replaced by $f_{n+1}(x, Z_i)$, the iteration would be a contraction and convergence would be guaranteed. However, such a procedure would require an iterative calculation for each new n, which would greatly increase computational time. And although the iteration defined by equation (28) is not generally a contraction, the procedure always seems to converge in practice, usually requiring about 20 iterations for the maximum distance between successive functions to be less than 10^{-5}. Furthermore, the inversion of a function on a grid is straightforwardly accomplished by reversing the roles of the grid and the function, and once again using cubic splines for interpolation, now using the non-uniformly spaced grid generated by the function values.

Figure 4 shows a set of price functions for the parameter values (0·22, −0·62, 0·024, 0·89). The horizontal axis shows the availability, here confined to a 20-point grid from −10 to 40, and the vertical axis the resulting price. At very low availability, price lies along the linear price function $P(x) = 0{\cdot}22 - 0{\cdot}62x$, and as availability increases, price falls along the $N = 10$ branches, one for each harvest. Because the autoregressive parameter here is positive (0.89),

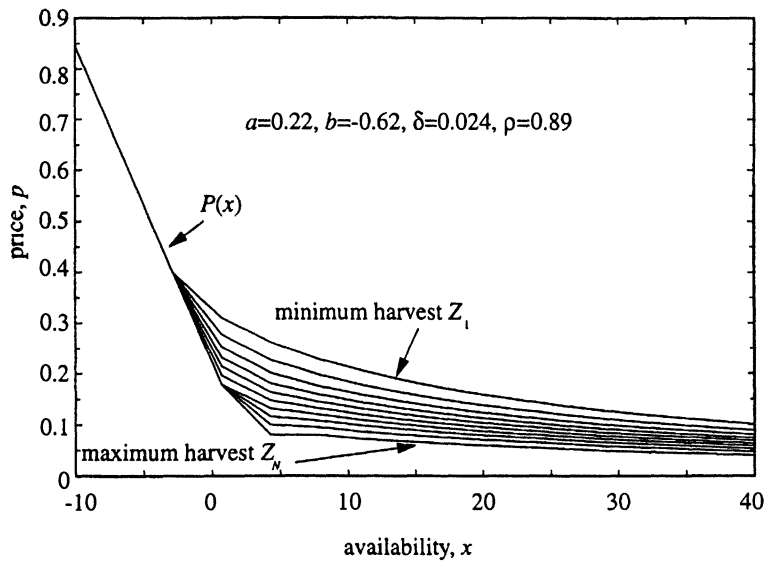

Figure 4. Price as a function of availability and harvest

larger harvests mean the expectation of larger harvests in the future and hence lower prices. If the autoregressive parameter were zero, which would be the i.i.d. case, the 10 branches of the price function would coincide, and there would be a single monotonic decreasing function linking price to availability.

3.3. Computing Invariant Distributions

The computation of the price function enables us to characterize the way in which the state variables evolve over time. The harvest z_t evolves according to the discrete approximation of the AR(1) described above, and does so in a way that is unaffected by the amount on hand x_t, while x_t evolves according to equation (10), whereby the new harvests are added on to any inventories inherited from the previous period. Starting from some initial position, the equations of evolution could be used to simulate the process, drawing new harvests in each period, and the simulations used to characterize the limiting invariant joint distribution of the harvests and the amounts on hand. The empirical invariant distributions are useful to characterize the behaviour of the model for any given set of parameter values, and thus to compare characteristics of the estimated price process with actual prices. Indeed, in the analysis of the i.i.d. model in Deaton and Laroque (1992), one of our major concerns was the inability of our simulations to reproduce the degree of positive autocorrelation in the data, and we shall want to make the same comparison using the estimated parameters. When we come to the case of autocorrelated harvests, we shall make even more essential use of the invariant distribution, since our estimation procedure will require its computation as part of the calculations.

Simulation is only one way of obtaining the limiting distribution, and for these models, it turns out to be relatively slow and inaccurate. The nonlinearities in the speculative storage model will occasionally produce sharp spikes in the price series, and because these are rare, large numbers of simulations are required to obtain accurate assessment of the limiting distribution. A more rapid and accurate procedure is to approximate the process (x_t, z_t) by a discrete Markov chain so that its evolution can be characterized by a transition matrix. This transition matrix has a unit eigenvalue whose eigenvector is the invariant distribution. The calculations are rather different in the i.i.d. and autoregressive cases, and we begin with the former.

Rewrite the transition equation (10) in the i.i.d. case in the form

$$x_{t+1} = (1-\delta)(x_t - P^{-1}(f(x_t))) + z_{t+1} = w(x_t) + z_{t+1} \tag{29}$$

thus defining the quantity $w(x_t)$. We already have a grid of points for x_t, X, and we calculate the transition probability from X_j to X_i from

$$T_{ij} = Pr(X_i + \Delta/2 \geq x_{t+1} \geq X_i - \Delta/2 \mid x_t = X_j) \tag{30}$$

where Δ is the distance between points on the grid X. Substituting from equation (29), and treating the distribution of z as if it were unit normal, we have

$$T_{ij} = \Phi(X_i + \Delta/2 - w(X_j)) - \Phi(X_i - \Delta/2 - w(X_j)) \tag{31}$$

Given the matrix T, the probabilities of each of the states are updated by

$$\pi_{it+1} = \sum_j T_{ij}\pi_{jt} \tag{32}$$

so that the invariant distribution π is the normalized eigenvector of T corresponding to the unit eigenvalue. The vector π may be conveniently obtained by solving the linear equations

$$\begin{pmatrix} T-\mathrm{I} & e \\ e' & 0 \end{pmatrix}\begin{pmatrix} \pi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{33}$$

where e is an M-vector of units.

The limiting distributions can be used to calculate summary statistics for the price in the limiting distribution. In particular, the mean, variance, and autocorrelation of prices are given by

$$\mu_p = \sum_{k=1}^{M} \pi_k f(X_k), \qquad \sigma_p^2 = \sum_{k=1}^{M} \pi_k f^2(X_k) - \mu_p^2 \tag{34}$$

and

$$\rho_p = \sigma_p^{-2} \sum_{i=1}^{M}\sum_{j=1}^{M} (f(X_j) - \mu_p)(f(X_i) - \mu_p) T_{ij} \pi_j \tag{35}$$

The autoregressive case follows the same general principles, although we now have to keep track of two state variables instead of one. Using the fact that z_t follows an autoregression, we can rewrite equation (10) as

$$x_{t+1} = (1-\delta)(x_t - P^{-1}(f(x_t, z_t))) + \rho z_t + \varepsilon_{t+1} = g(x_t, z_t) + \varepsilon_{t+1} \tag{36}$$

where ε_{t+1} is the innovation in the harvest process. For reference, we rewrite equation (7) the autoregression for z_t as

$$z_{t+1} = \rho z_t + \varepsilon_{t+1} \tag{37}$$

We suppose that at time t, the amount on hand and the harvest (x_t, z_t) are at the grid point (X_k, Z_l) and we are interested in the transition probability

$$T_{ijkl} = Pr((X_i + \delta/2 \geq x_{t+1} \geq X_i - \delta/2) \cap (\theta_{j+1} \geq z_{t+1} \geq \theta_j) \mid x_t = X_k, z_t = Z_l) \tag{38}$$

Substituting from equations (36) and (37), equation (38) becomes

$$T_{ijkl} = Pr((X_i + \delta/2 - g(X_k, Z_l) \geq \varepsilon_{t+1} \geq X_i - \delta/2 - g(X_k, Z_l)) \cap (\theta_{j+1} - \rho Z_l \geq \varepsilon_{t+1} \geq \theta_j - \rho Z_l)) \tag{39}$$

Note that this probability will only be non-zero if ε_{t+1} is simultaneously in both intervals. Equation (39) is evaluated using the standard normal distribution for the innovations. It is then converted into a two-dimensional matrix by stacking into an $MN \times MN$ transition matrix, the first M rows of which refer to arrival at the M-values of X and at Z_1, the second M rows to arrival at the M-values of X and at Z_2 and so on. The limiting joint distribution of (x_t, z_t) is then calculated using a larger dimension version of equation (33). This MN vector is most conveniently rearranged into a M by N matrix containing the grid point frequencies that approximate the joint distribution.

Figure 5 illustrates the calculated invariant distribution for cotton in the autoregressive case. The figure shows the strong positive correlation between the harvest and the amount on hand in the limit distribution. Since storage cannot be negative, the distribution is above the 45-degree line in the horizontal plane, tending to be quite concentrated near the line when harvests are low, but stretching out when harvests are high.

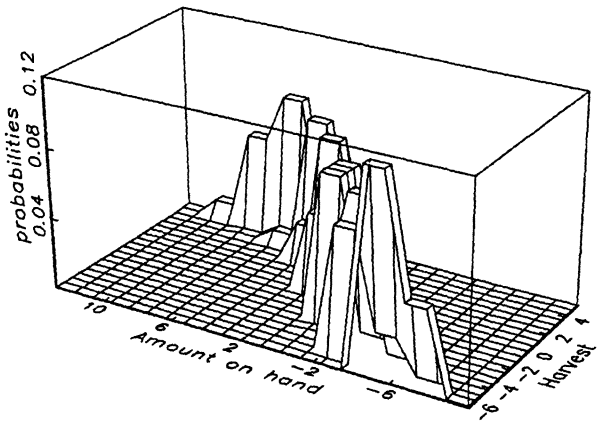

Figure 5. Invariant probabilities

4. ESTIMATION AND RESULTS OF THE I.I.D. CASE

4.1. Estimation

The case where harvests are independently and identically distributed is relatively straightforward to estimate, and we develop it first. It allows us to dispose of a number of preliminaries, and the results from the i.i.d. estimation will provide the motivation for tackling the AR case in the next section.

When the harvests in each period are independent of previous harvests, the price function (26) has only a single argument and collapses to

$$f(x) = \max\left\{N^{-1}\sum_{j=1}^{N} f[Z_j + (1-\delta)(x - P^{-1}[f(x)])], P(x)\right\} \tag{40}$$

For any given set of data p_t, $t = 1, \ldots, T$, and parameters δ, a, and b—the last two the parameters of the price function—$f(x)$ is calculated to satisfy equation (40) using the iterative methods discussed in the previous section. The pseudo-likelihood function is then calculated as follows. Given any price p_t, for t up to $T-1$, we use equation (16) to calculate the one-period-ahead expectations and variances according to

$$m(p_t) = E(p_{t+1}|p_t) = N^{-1}\sum_{i=1}^{N} f[(1-\delta)(f^{-1}(p_t) - P^{-1}(p_t)) + Z_i] \tag{41}$$

and

$$s(p_t) = V(p_{t+1}|p_t) = N^{-1}\sum_{i=1}^{N} f^2[(1-\delta)(f^{-1}(p_t) - P^{-1}(p_t)) + Z_i] - m^2(p_t) \tag{43}$$

These two expressions are then inserted into the pseudo-likelihood function (12), thus establishing the pseudo-likelihood as a function of the parameters (and the data.) This function is then maximized following some standard procedure that uses numerical derivatives.

The evaluation of $f^{-1}(p_t)$ in equations (41) and (42) calls for some comment. The calculations discussed in the previous section yield the function $f(x)$ at each point of the grid X. Corresponding to each point on X there is a price, and these prices $f(X)$ can be regarded as a grid—albeit not equally spaced—for the function $f^{-1}(p)$, with corresponding values X. The actual prices p_t are located in this grid and the associated values of the state variable x are calculated by cubic spline interpolation. Figures 6–8 show how these procedures work for the case of cotton. In Figure 6, the grid points for X, the amount on hand, are marked on the vertical axis; these points are equally spaced. Price is marked along the horizontal axis, and the smooth curve shows our approximation to $f^{-1}(p)$ computed by using the marked price points $f(X)$ as grid points and fitting cubic splines. Because the transform is nonlinear, the price grid points are more heavily concentrated among low prices. The smooth curve makes no explicit

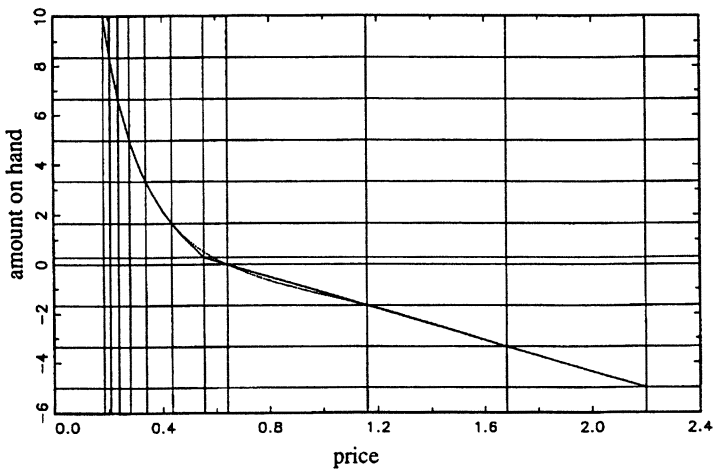

Figure 6. Cotton: i.i.d. case; function $f^{-1}(p)$

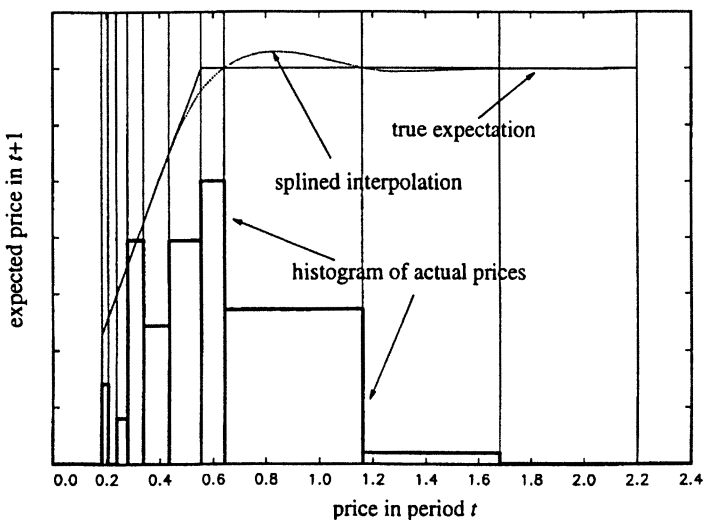

Figure 7. Cotton: i.i.d. case; conditional expectation $E(p_{t+1} \mid p_t)$

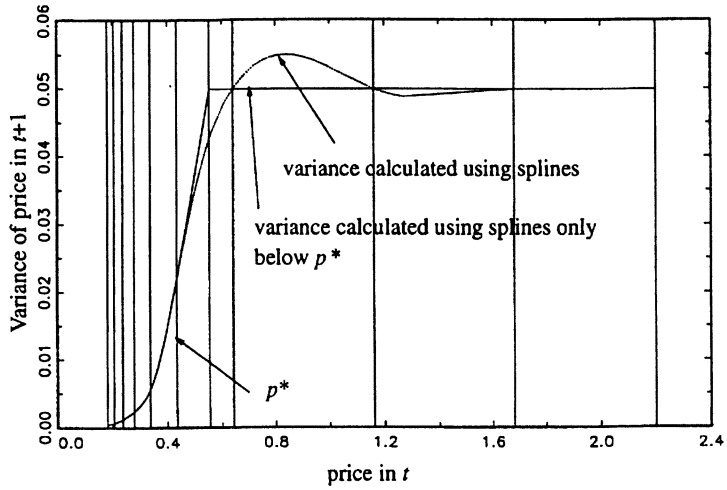

Figure 8. Cotton: i.i.d. case; conditional variance $V(p_{t+1} \mid p_t)$

allowance for the kink point at $p = p^*$, and the second curve shows what happens if we fit the spline only for $p < p^*$ and use the fact that for $p > p^*$, $x = (p - a)/b$. In this case, the smooth curve is a good approximation. The functions in this figure are simply the inverse of the functions in Figure 4 in the i.i.d. case where the branches of Figure 4 collapse to a single curve.

Figure 7 shows the distribution of actual prices along the price grid, as well as two different approximations to the conditional expectation. The kinked line is the one-period-ahead conditional expectation of price using the piecewise linear form (19) with p^* calculated from the price function using equation (17), while the smooth line is calculated directly from equation (41) using the smooth versions of $f(x)$ and $f^{-1}(p)$. The heavy horizontal lines running between the boundaries of the price grid show the relative frequencies of the prices in the data; the heaviest concentration of prices is a little below $p = p^*$. Figure 8 shows the comparable calculations for the one-period-ahead variance, and once again we see that the smooth curves are not always very accurate approximations, especially over areas where the spline knots are far apart. The estimates for the i.i.d. case reported below use the smooth curves with the splines smoothing over the kink points. Although this is less accurate than the piecewise linear form, the non-differentiability in the latter slows the convergence of our numerical algorithm.

A potential problem with the price grids is that there is no guarantee that, for any given parameter values, one or more of the actual data points will not lie outside the range $[f(X_M), f(X_1)]$ in which case we have no means of calculating $f^{-1}(p)$ in equations (41) and (42). In these cases, we have found no better solution than stopping the algorithm, resetting the grid, and beginning again. Changing the grid within the estimation procedure is hazardous because grid changes alter likelihood values even without a change in parameter values, so that control of the algorithm becomes difficult. One possibility would be to add points to a pre-existing grid, leaving previous grid points intact, and although this should not alter the likelihood by much, the additional points would affect the cubic splines for the original points, and the likelihood evaluations would be affected. In practice, we were guided by the maxima and minima discussed in the previous section, but proceeded by trial and error. Table I lists the lower and upper limits for the grids that were used for each of the commodities; we make no claims that persistent search might not generate tighter bounds. The table also shows the number

Table I. Grid limits and sizes for each commodity

	Lower limit	Upper limit	Grid size
Cocoa	−30	15	20
Coffee	−6	20	10
Copper	−10	50	20
Cotton	−5	10	10
Jute	−5	15	10
Maize	−5	30	10
Palm oil	−10	40	20
Rice	−5	10	10
Sugar	−5	10	10
Tea	−5	15	10
Tin	−5	40	20
Wheat	−3	50	20

of grid points for each good. In all cases, we began with ten, and increased it if the algorithm showed difficulty in finding a flat maximum of the likelihood, the idea being that a finer grid would allow a better approximation to the likelihood function and (perhaps) allow easier convergence. The fineness of the grid over x varied from commodity to commodity, but 10 or 20 points were generally sufficient to generate a well-defined maximum.

Two of the parameters, b and δ, must satisfy restrictions for the model to make sense. In particular, b must be negative and $\delta > -r$ for the model to be coherent and to guarantee that there exists a non-negative price function $f(x)$. It is not possible to estimate without these restrictions and check whether they are satisfied *ex post*; the calculations cannot be carried through for values outside the correct ranges. To guarantee the inequalities, we estimate the three parameter vector θ defined by (recall that r is set at 0·05):

$$a = \theta_1 \qquad b = -e^{\theta_2} \qquad \delta = -0{\cdot}05 + e^{\theta_3} \tag{43}$$

4.2. Evaluation of Results

Once the problem was set up as described, we had no great difficulty in obtaining parameter estimates using a form of the Berndt *et al.* (1974) algorithm with numerical derivatives evaluated by small perturbations. Although the algorithm is based on the equality in expectation of the Hessian of the log likelihood function and the outer product of its gradient vector, an equality that does not hold in the case of pseudomaximum likelihood, the method appears to work satisfactorily in our case. We used the gradient of the objective function as a convergence criterion, and the calculations were stopped when the elasticity of the criterion with respect to each parameter was less than 1.E–04.

We ran one illustrative Monte Carlo experiment in order to check that the procedures were capable of generating sensible results. With a set to 0·20, b to −0·15, and δ to 0·12, values that are reasonably representative of those that we estimate, we computed the price function and used it to generate 100 samples of 100 price observations each. These were then fed into the estimation procedure, using the true values as starting values, and parameter estimates obtained. The algorithm was stopped when convergence was obtained, 88 cases, after 80 iterations if convergence had not occurred, seven cases, or when no improvement in the pseudo-likelihood could be located along the line of steepest ascent, one case. In the remaining four cases, the

generated prices fell off the price grid, and the estimation was aborted; these cases are excluded from the following results. The mean of the parameter estimates so obtained was 0·1983 for a, −0·1468 for b, and 0·1245 for δ. Over the 96 replications, the standard deviations were 0·01507, 0·01888, and 0·03905, respectively, as compared with the mean of the calculated robust standard errors of 0·01538, 0·01940, and 0·04000. While these very limited results are hardly an adequate test of the performance of the estimator, they do not suggest any obvious problems.

In the practical applications, we had some concern about the presence of multiple maxima. For three of the commodities, maize, rice, and sugar, we ran further Monte Carlo experiments in an attempt to assess the danger. For each good, we ran 100 trials starting from perturbations of the final values of the θ parameters obtained by multiplying the θ's by random numbers drawn from a lognormal distribution, so that the standard deviation of the log perturbation was 5%. For maize, 99 of the 100 experiments converged to the same values, and one converged to different estimates with a log pseudo-likelihood of 32·051, as opposed to 32·056 for the other 99 simulations. This is the only case of a multiple maximum we have found, in the maize experiments, or in those for the other two commodities. In these experiments, the modal number of iterations to convergence was 17, with a minimum of nine and a maximum of 24. Of the 100 experiments for rice, one failed to produce estimates because of an inadequately large grid, and the other 99 converged to the same estimates. In this case, more iterations were required, with a minimum of 28, a maximum of 60, and a mode of 40; 79 of the 99 cases converged in 35–44 iterations. For sugar, there were seven cases where the starting values led to grids that were too small, while the other 93 converged to the same values. The minimum number of iterations was 11, the maximum 26, and the mode 18.

Table II shows the estimates and standard errors for a, b, and δ as well as the predicted and actual autocorrelations. These estimates and their implications are discussed in some detail in Deaton and Laroque (1996), but the major issue for this paper is the failure of the model to account for the autocorrelation in any of the twelve commodities. In Deaton and Laroque (1992) our simulations failed to generate autocorrelations as high as those in the data, and the calculations in the final column of Table II confirm that this result did not reflect our inability to choose the right parameters in the simulations, but is a general feature of the model. Figure 9,

Table II. Parameter estimates and predicted and actual autocorrelations

							Autocorrelation	
	a	s.e.	b	s.e.	δ	s.e.	Model	Actual
Cocoa	0·16	(0·01)	−0·22	(0·03)	0·12	(0·04)	0·30	0·83
Coffee	0·26	(0·02)	−0·16	(0·03)	0·14	(0·02)	0·24	0·80
Copper	0·54	(0·04)	−0·33	(0·05)	0·07	(0·02)	0·39	0·84
Cotton	0·64	(0·04)	−0·31	(0·04)	0·17	(0·03)	0·19	0·88
Jute	0·57	(0·03)	−0·36	(0·06)	0·10	(0·05)	0·30	0·73
Maize	0·63	(0·04)	−0·64	(0·15)	0·06	(0·03)	0·36	0·76
Palm oil	0·46	(0·05)	−0·43	(0·06)	0·06	(0·03)	0·42	0·73
Rice	0·60	(0·03)	−0·34	(0·03)	0·15	(0·04)	0·22	0·83
Sugar	0·64	(0·05)	−0·63	(0·06)	0·18	(0·03)	0·26	0·62
Tea	0·48	(0·02)	−0·21	(0·02)	0·12	(0·03)	0·23	0·78
Tin	0·26	(0·04)	−0·17	(0·05)	0·15	(0·05)	0·26	0·90
Wheat	0·72	(0·04)	−0·39	(0·03)	0·13	(0·03)	0·26	0·86

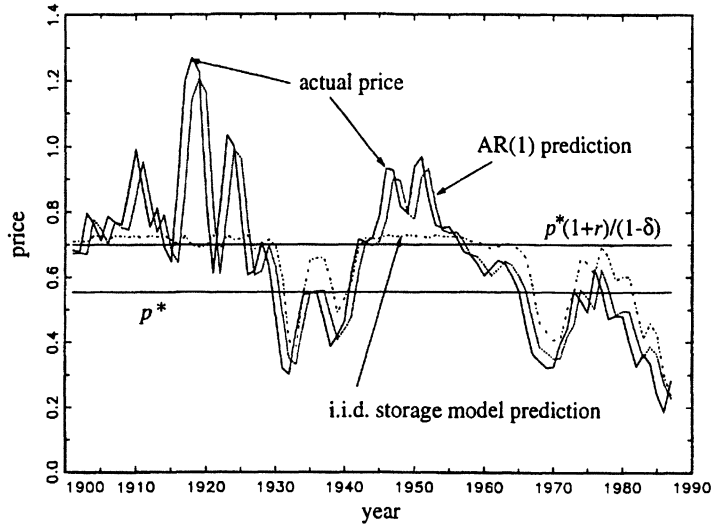

Figure 9. Cotton: actual and predicted prices

again for cotton, shows why it is that the model fails to fit the data. The graph shows the data points together with the one-period-ahead predictions from the storage model and from a simple AR(1), which is the process that the price would follow if the harvest were an AR(1) and there were no storage, as would occur, for example, if δ were close to unity. The AR(1) is not a very good model, but it does better than the storage model, largely because the storage model cannot account for autocorrelation in prices when price is above the cutoff p^*. At this point, stocks are not being held, and next period's price is unaffected by the current price (see equation (19)). But this is inconsistent with the behaviour of actual prices, which appear to be autocorrelated even when prices are high.

This failure of the model is our main motivation for turning to the case where the harvests are autocorrelated. In such cases, there will be autocorrelation in prices even in the absence of storage, something that appears to be required by the data. Although speculative storage is capable of generating some autocorrelation in prices even when the underlying shocks are i.i.d., the predicted autocorrelation is far short of that in the actual prices. Since it is plausible that the underlying shocks are themselves autocorrelated, the obvious next question is whether autocorrelated shocks together with speculative storage can account for the data. The estimation of such a model is the topic of the next section.

5. ESTIMATION WITH AUTOREGRESSIVE HARVESTS

5.1. Procedures

As in the i.i.d. case, we start from the price function and combine it with the data so as to generate one-period-ahead expectations and variances that can be inserted into the pseudo-likelihood function. The major difference in the autoregressive case is that the price function has two arguments, the amount on hand and the harvest itself, so that it is no longer possible to infer the two state variables from the single observed price. However, for each value of the harvest z, we can look for an amount on hand leading to the observed price through inversion of the price

function. Hence, using equation (26), and the T prices p_t, we attempt to solve

$$x_{ti} = f^{-1}(p_t, Z_i) \tag{44}$$

for $t = 1, \ldots, T-1$, and $i = 1, \ldots, N$. Conditional on Z_i, $f^{-1}(p_t, Z_i)$ is monotonic in p_t, so that high prices may require an x lower than the bottom of the grid X_1 and low prices an x higher than the top of the grid X_M. In these cases, we use X_1 or X_M respectively for x_{ti} in equation (44). We shall return to these and related difficulties below.

Excluding these boundary cases, this procedure yields up to N possible x's consistent with each price, one for each possible value of the harvest. Without observing the latter, all x's are consistent with the price observation. In constructing the one-period-ahead expectations and variances, we construct an expectation and variance for each of the N pairs (x_{ti}, Z_i), i.e. expectations and variances conditional not only on the current price but also on a specific harvest, and then weight the results so as to condition only on price. We describe first the construction of the moments conditional on the harvests, and then the weighting, turning finally to the various practical and theoretical difficulties, including what happens when equation (44) has no solution.

The means and variances in the i.i.d. case (equations (41) and (42)) can readily be extended to the AR case conditional on the harvest, i.e.

$$\begin{aligned} m(p_t, Z_i) &= E(p_{t+1} \mid p_t, z_t = Z_i) \\ &= \sum_{j=1}^{N} T_{ji}(\rho) f[((1-\delta)(f^{-1}(p_t, Z_i) - P^{-1}(p_t)) + Z_j), Z_j] \end{aligned} \tag{45}$$

and the comparable conditional variance

$$\begin{aligned} s(p_t, Z_i) &= V(p_{t+1} \mid p_t, Z_i) \\ &= \sum_{j=1}^{N} T_{ji}(\rho) f^2[((1-\delta)(f^{-1}(p_t, Z_i) - P^{-1}(p_t)) + Z_j), Z_j] - m^2(p_t, Z_i) \end{aligned} \tag{46}$$

For the case of cotton, Figures 10–12 show how this procedure works. Figure 10, which is essentially the inverse of Figure 4, shows the calculated inverse price functions (44) for each of the harvests on the grid Z. The harvest autocorrelation coefficient built into this figure is positive, so that, conditional on the amount on hand, higher harvests mean lower prices. In consequence, a given price requires a lower amount on hand at higher harvests. The function $f(x, Z)$ is calculated for all x in X and for all Z, so that the solution of equation (44) may yield an x_{ti} smaller than Z_i. However, along any trajectory generated by the model, the amount on hand cannot be less than the harvest. The solid line at the bottom of the figure shows the relationship $p = f(Z, Z)$ plotted with Z on the vertical axis and p on the horizontal. Since price is monotonically decreasing in the amount on hand, this relationship shows the smallest amount on hand consistent with any given price and harvest combination. As the graph shows, this amount may be slightly larger than the bottom part of the curves, especially for large harvests and low prices.

Figures 11 and 12, again for cotton, show the calculated values of equations (45) and (46). Figure 11 shows that, conditional on a given value of the harvest, the expected price is a piecewise linear function of the current price (see equation (20)). Lower harvests generate higher expected prices when harvests are positively autocorrelated. The conditional variances in Figure 12 are not something that we have been able to characterize *a priori*, except that they should be monotonically increasing for a given z. The calculations in this case suggest a good

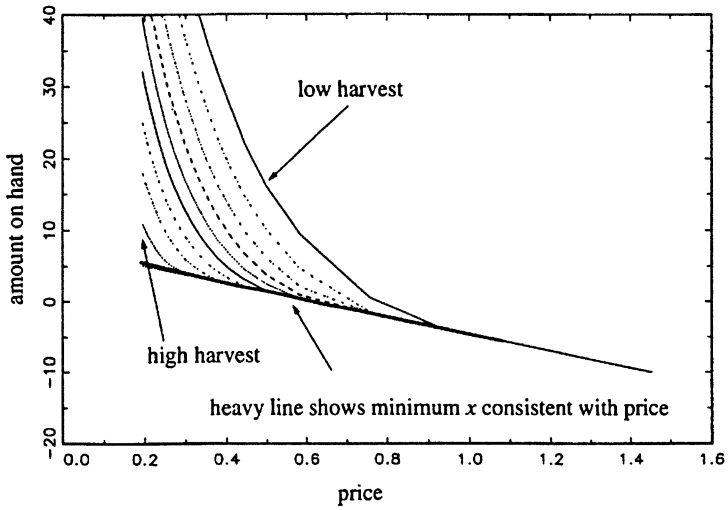

Figure 10. Inverse price functions and harvests

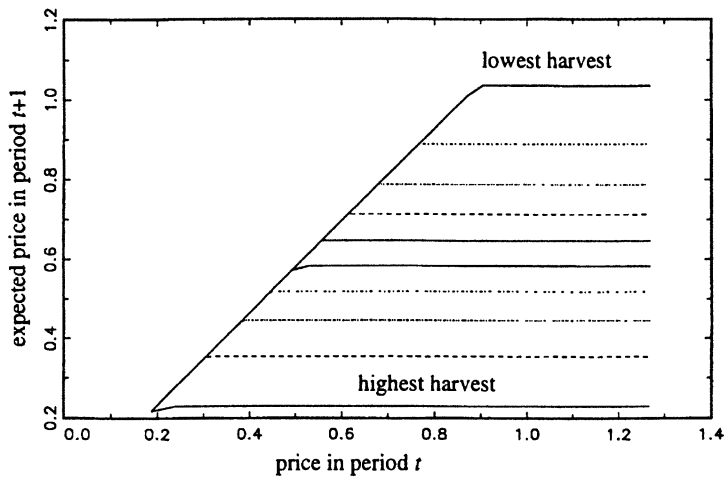

Figure 11. One-period-ahead expectations of price, conditional on p and z

deal of complexity. The variances appear to increase with the harvest at low prices. When prices are high today, there are no carried-over inventories; the conditional variances are constant and reflect only the variability of the next period's harvest as transformed into next period's price. If the harvests were an AR(1) with continuous normal innovations, we would therefore expect the variance of prices to decrease with the current harvest given that the function f is convex. We interpret the non-monotonic behaviour shown in the figure as a consequence of our discrete parametrization of the harvest, whereby the conditional variances of z_t have the shapes shown in Figure 2.

To convert equations (45) and (46) into expectations that can be used in the estimation procedure, we must integrate out the unobservable Z_i, for which we require probabilities conditional on some set of observations available to the econometrician. There are several ways

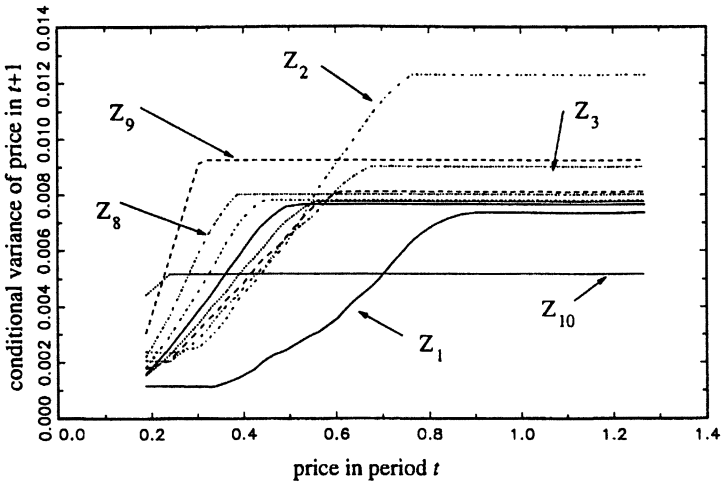

Figure 12. Variance of price in $t+1$ conditional on price and harvest$t$

of doing this. One is to start from some arbitrary conditional density for the harvests for p_1, and to update it period by period, so as to yield the density of Z_t conditional on all prices back to p_1. This method has the advantage of using all the available information. We plan to pursue this option in later work, but we have not done so yet. Instead, we adopt a (presumably) less efficient strategy where we condition only on current price p_t. Using Bayes' rule and equation (44) gives for all the x's and Z's consistent with p_t,

$$\gamma_{it} = \Pr(Z_i \mid p_t) = \frac{\Pr[x_t = f^{-1}(p_t, Z_i), z_t = Z_i]}{\sum_{j=1}^{N} \Pr(x_t = f^{-1}(p_t, Z_j), z_t = Z_j]} \tag{47}$$

where the probabilities on the right-hand side are interpolated—using cubic splines as always—from the invariant distribution of (x_t, z_t) which is calculated for the grid points $X \times Z$ as described in Section 3.3.

The conditional expectation $E(p_{t+1} \mid p_t)$ is then computed according to

$$m(p_t) = E(p_{t+1} \mid p_t) = \sum_{j=1}^{N} \gamma_{jt} m(p_t, Z_j) \tag{48}$$

The conditional variance contains two components, the variance *within* harvests, which is the weighted average of the conditional variances in equation (47), and the variance *between* harvests, a component that would be zero if the current harvest were known. Hence

$$s(p_t) = s_w(p_t) + s_b(p_t) = \sum_{j=1}^{N} \gamma_{jt} s(p_t, Z_j) + \sum_{j=1}^{N} \gamma_{jt} [m(p_t, Z_j) - m(p_t)]^2 \tag{49}$$

The implementation of equations (47)–(49) presented considerable difficulties in practice. One problem occurs when we encounter a price p_t in the data for which we are unable to find values of $x \in [X_1, X_M]$ and $z = Z_i$, for some i such that $p_t = f(x, z)$. This can happen because (1) although according to the true model there exists a pair (x, z) that generates the observed price,

the grids X and Z are insufficiently large to contain it (recall that the grid Z contains N points that replace a normal distribution with infinite support), or (2) there is no combination of x and z that generates p_t so that the price is inconsistent with the model. Even when we can solve equation (44), there is no guarantee that the solutions (x, z) lie within the support of the invariant distribution, or at least within the support of the invariant distribution as calculated. Once again, the calculation of the conditional probabilities γ_{kt} will fail.

Case (1) can be solved in principle by enlarging either X or Z (or both), although our procedure of tying Z to the normal distribution makes it difficult to generate very large positive or negative values. However, there are also practical considerations that set limits to grid sizes. When the length of the grid is large, the cubic spline interpolation will become increasingly inaccurate—and possibly non-monotonic—unless additional grid points are added, i.e. unless M is increased. But increasing either M or N—the number of harvests—adds to computational time and can ultimately threaten the feasibility of the estimation. For example, the calculation of the invariant distribution requires solution of a system of MN equations, which is a major contribution to computing time. Note also that the distribution of commodity prices is strongly positively skewed, so that it will often be possible to find a short grid with accurate interpolation that accounts for all but a few prices, even though covering all the prices would require a very large grid. For these reasons, we do not always choose to extend the grid.

Case (2), or when the observed price is outside the invariant distribution, seems most appropriately dealt with by discarding the model since it cannot account for the data, at least in its steady state. But this is to treat the parameterization too seriously, and would be comparable (for example) to discarding a regression with normal errors because one residual is more than three standard errors from zero. In the current case, we have restricted the harvests to a discrete approximation to the normal, so that, when $N = 10$ the lowest harvest is 1·76 standard deviations below the mean, so that the model would have to be rejected whenever $p_t > f(-1{\cdot}76, -1{\cdot}76)$ the imposition of which would treat our parametrization more seriously than is appropriate.

In the light of these arguments, our operating procedure is as follows. When an observed price is larger than the maximum possible price $f(Z_1, Z_1)$ or less than the lowest possible price $f(X_M, Z_N)$, we place all of the weight on Z_1 or Z_N respectively. Hence

$$\begin{aligned} \gamma_{it} &= \delta_{i1} \qquad \text{if } p_t \geq f(Z_1, Z_1) \\ \gamma_{it} &= \delta_{iN} \qquad \text{if } p_t \leq f(X_M, Z_N) \end{aligned} \tag{50}$$

and the γ_{it} so extended are used to calculate the conditional moments and variances in equations (48) and (49). Equation (50) effectively assumes that, when the price is either small or large enough, we know the harvest—it is either the maximum or minimum—so that only one each of the conditional moments or variances (46) and (47) is relevant. As a result, the $s_b(p_t)$ term in (49) is zero; there is no contribution to the variance from the econometrician's uncertainty about the harvest.

There is another implication of equation (50) that is less appropriate. Suppose that $p_t > f(Z_1, Z_1)$ so that all the weight in the conditional probability γ_{it} is on the first element. According to equation (47), we have $m(p_t) = m(p_t, Z_1)$ a function that is constant for $p \geq p^*(Z_1)$ (see equation (20) above). In consequence, when the price is high enough, the one-period-ahead conditional expectation of price will be constant. Our intuition suggests—but we have no proof—that, when price is sufficiently high, the conditional distribution of the state variable will be such that there are stocks neither in the current period nor in the next period, in which case price, like the harvest itself, will be autoregressive and $m(p_t) = (1 - \rho)a + \rho p_t$, the

derivative of which with respect to p_t is ρ, which in the calculations here is typically closer to 1 than 0. While it would be possible to replace $m(p_t, Z_1)$ in equation (48) by a similar expression for prices above some critical level, this turns out not to be a satisfactory solution. When we look at $m(p_t)$ from equation (48) in the range where p_t is high enough for the conditional probabilities to be weighted towards the low harvests, but not yet concentrated on the lowest, the slope is usually close to ρ, so that, on the basis of this observation, we extrapolate $m(p_t)$ from the last 'uncontaminated' values with slope ρ. The critical point for p_t is set at the stocking-out point corresponding to the second lowest harvest $p^*(Z_2)$ so that $m(p_t)$ in equation (48)—but not in equation (49)—is replaced by

$$\tilde{m}(p_t) = m(p_t)1(p_t \leq p^*(Z_2)) + [m(p^*(Z_2)) + \rho(p - p^*(Z_2))]1(p_t \geq p^*(Z_2)) \qquad (51)$$

Equation (51) is an *ad hoc* fix without theoretical justification. However, it will be exactly correct in the case where there is no storage so that the price process, like the harvest process, is an AR(1). Perhaps because this model is close to what we estimate for the full storage model, equation (51) appears to provide a sensible approximation, something that we have examined extensively through plots of both m and $\tilde{m}$.

Figure 13 shows the calculated conditional expectation, again for the case of cotton and using the best-fitting parameter values. Given that the autocorrelation coefficient is so large, 0·92, this curve is effectively that of an AR(1), particularly once equation (51) is used to replace the directly calculated dotted line by the solid line. Figure 14 shows the two components of the variance s_w and s_b from equation (49). The within variance, labelled 'expectation of variance', is the component of the variance associated with uncertainty about the future harvest, uncertainty that is common both to the farmers and to the econometrician. The between variance, labelled 'variance of expectation,' comes from the fact that the current harvest is unknown to the modeller. Since we have no analytical results about the properties of either of these variances, or indeed of their sum in Figure 15, it is difficult to know whether the shapes in the figures are genuine or reflect the various approximations in the calculations. Given that the estimates of ρ are so high, and given the behaviour of the conditional variance of the harvest for high ρ in Figure 2, the discrete approximation to the continuous autoregression is a possible source of the behaviour depicted in Figures 14 and 15.

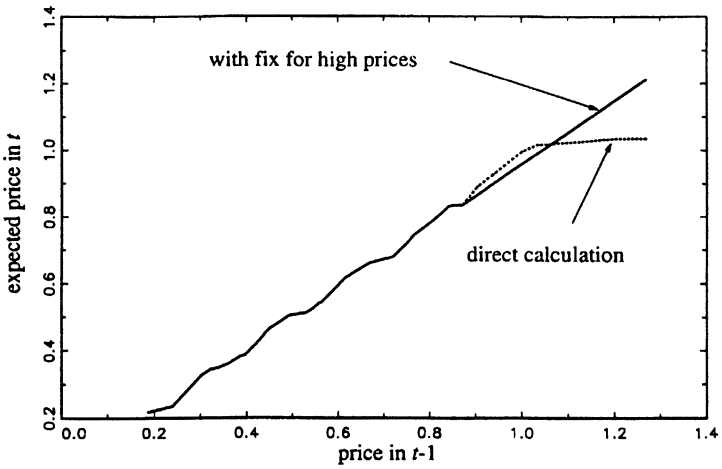

Figure 13. Expectation of price in t conditional on price in $t-1$

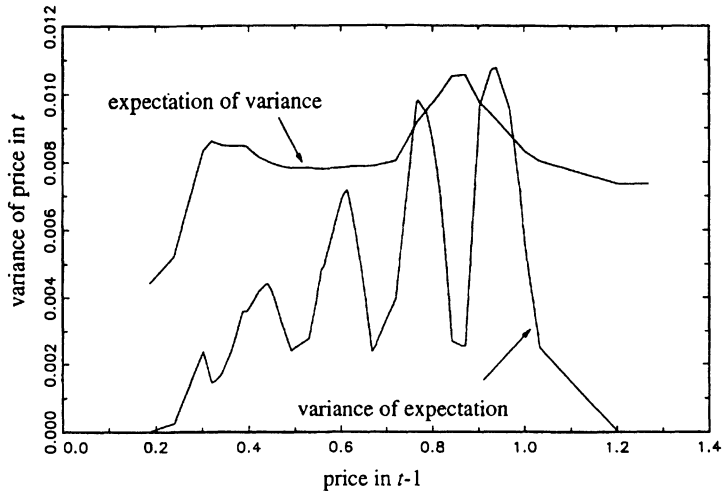

Figure 14. Components of conditional variance

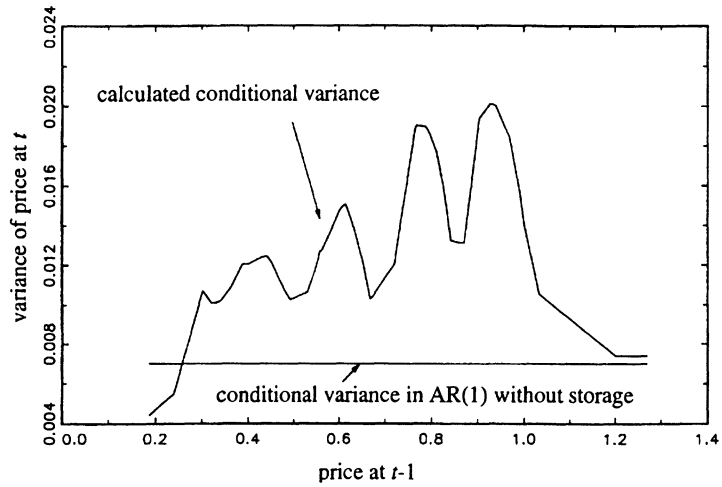

Figure 15. Variance of price at t conditional on price at $t-1$

5.2. Results

Once again it seemed wise to begin with a limited Monte Carlo experiment, if only as a check of the code. In fact, the results provided a good deal more, and gave insights into the operational characteristics of the estimation under the best possible circumstances, when the model is correct and when the estimation can be started from the true values of the parameters. We used the same parameters as for the i.i.d. case (0·2, −0·15, 0·12), supplemented with an autoregressive parameter ρ of 0·5. We generated 100 price samples of 100 observations each. Of the 100 experiments, 59 reached convergence within 50 iterations, 17 had not converged after 50 iterations, and the remaining 24 were terminated when the algorithm could not find a higher pseudo-likelihood value along the direction of steepest ascent. Over all experiments, the

average of the estimates was 0·2064, −0·1557, 0·1370, and 0·4070, with sample standard deviations of 0·0351, 0·0456, 0·0565, and 0·2502. The autoregressive parameter appears to be the most difficult to estimate well. It is not only subject to a large amount of sampling error, but the mean estimate is almost four standard errors below the true value of 0·5. Because so many of the experiments did not reach convergence, the estimated robust standard errors are not useful, sometimes producing absurdly high and low values. If we condition on the 59 cases where convergence was obtained, the average calculated standard errors are 0·0306, 0·0479, 0·0446, and 0·2421, which correspond reasonably well to the standard deviations for the full set of 100 experiments. The means and standard deviations of the parameter estimates are not significantly (or substantively) different depending on whether or not the algorithm converged, a result that may well be sensitive to our having begun the iterations from the true values of the parameters.

The i.i.d. model of Section 4 is a special case of the autoregressive model with storage (when $\rho = 0$) as is the standard AR(1) without storage (δ large), although note that if price follows an AR(1) with normal errors, negative prices are not excluded *a priori*, whereas, in our model, negative prices are ruled out by the possibility of storage. In consequence of the nesting, we can use the code for the full model together with the estimated parameters of the other models to recalculate pseudo-likelihoods for the i.i.d. with storage and AR(1) without storage models. The results will not be exactly the same because of the various approximations and choice of grids. The standard likelihood for the AR(1) without storage assumes a continuous normal distribution, and although the grid for the harvest is set using the same parameters as for the i.i.d. case, and as listed in Table I, the range of the harvests is larger the larger is ρ. The maximum harvest is no longer Z_N but $Z_N/\sqrt{1-\rho^2}$, and the grid for the amount on hand must accommodate this. Nevertheless, that the different calculations of the likelihoods should give similar answers is an important check on the credibility of our code.

Table III lists the pseudo log likelihood values for the i.i.d. model with and without storage, the AR(1) model with no storage, and the full model of this section, in which the harvests are AR(1) and storage is allowed. The results in the first panel are calculated by standard methods, and are the values of the log likelihood functions associated with (1) the supposition that the prices are i.i.d. normal, which would be the case if the harvests were i.i.d. normal and there were no storage, and (2) that prices follow a continuous AR(1) process with normal innovations, as would be the case if the harvests followed a similar process and there were no storage. These values are calculated excluding the first observation so that the same data points are included in all likelihood calculations. Given that the commodity with the least first-order autocorrelation (sugar) has an autocorrelation coefficient of 0.62, it is hardly surprising that the numbers in the second column are all much larger than those in the first.

The second panel shows the log pseudo-likelihood value from the parameter estimates in Table II and comes from the estimation procedure of Section 4. These values are all larger than those for the i.i.d. model without storage in the first column, but only for cocoa, maize, and sugar does the model do better than the simple AR(1). Although these two models are not nested, they are both special cases of the more general model, and the much simpler AR(1) typically does better.

The results in the final panel are calculated using the methods of this section. The first two columns report the log pseudo-likelihood values for the i.i.d. model with storage and the AR(1) model without storage. These are obtained using the parameters from the middle panel; together with $\rho = 0$, and from the second column of the first parameter with δ set to 0·9. Apart from the case of palm oil in the AR model, there is a reasonably close correspondence between the values from the different procedures in spite of the very different methodologies and

Table III. Log pseudo-likelihood values for various models

	i.i.d. model	AR model	i.i.d. storage model	AR harvest model with storage		
				i.i.d.	AR	Full model
Cocoa	71·9	124·1	125·2	123·6	117·6	141·6
Coffee	75·0	118·9	111·0	109·9	110·2	131·2
Copper	24·7	81·1	73·9	75·0	73·2	103·1
Cotton	6·76	74·2	29·8	33·9	69·6	79·0
Jute	18·9	50·2	44·8	42·0	47·2	53·3
Maize	−10·7	27·0	32·1	30·6	22·2	46·2
Palm oil	−5·76	27·6	22·2	24·1	17·3	58·9
Rice	6·92	61·0	26·0	22·0	55·9	61·0
Sugar	−48·6	−27·0	−10·7	−15·1	−28·9	−3·90
Tea	54·5	100·9	69·3	67·0	99·8	102·5
Tin	84·015	150·9	108·9	106·1	149·5	152·3
Wheat	−7·77	52·8	24·6	22·0	50·0	53·2

Notes: The first panel shows the likelihoods for (a) the price follows an i.i.d. normal process, and (b) the price follows an AR(1) process with i.i.d. normal innovations. The second panel shows the log pseudo-likelihood for the storage model with i.i.d. harvests using the estimation procedures discussed in Section 4. The final panel shows log pseudo-likelihood values for the procedures described in Section 5. The first two columns are the likelihood values associated with the parameters from the storage model with i.i.d. harvests and with the parameters from an AR(1) price process without storage; up to the approximation, these numbers should equal the numbers in columns (3) and (2) respectively. The final column is the highest value of the pseudo-likelihood that we have been able to find.

approximations. The final column in the last panel shows the highest values of the pseudo-likelihood that we have been able to obtain for each of the commodities. It is far from clear that these values are the maximum of the pseudo-likelihood function. The gradients are not zero at the listed values except for the single case of copper, and for one good, jute, the algorithm converged to a point with zero gradients that has a lower pseudo-likelihood than that listed. Given the non-differentiability of the calculated expectation and variance functions, as well as the various fixes discussed above, it is not to be expected that the maxima would be located at points of zero gradient. This fact deprives us of any useful standard errors for the parameter estimates.

The likelihood values in the last column of Table III are higher than those in the previous two columns, as they must be, and apart from tea, tin, and wheat, the increase in fit is a good deal more than would usually be regarded as justified from the addition of a single parameter. Table IV lists the corresponding parameter estimates together with the raw autocorrelation coefficient of prices, replicated from Table II. Several points deserve comment. For three of the commodities, copper, maize, and palm oil, for all three of which the AR model with storage shows a marked improvement over the others, our estimation procedure forces δ to (or close to) its lower limit of $-0{\cdot}05$, which is (minus) the value at which the real interest rate is set. In these cases, the *negative* deterioration of stocks compensates exactly for the positive interest costs, so that carrying costs are zero, and inventories can become indefinitely large. Our calculations are likely to be especially unreliable in this case, the invariant distribution is concentrated at the top of the grid for the amount on hand, as would be expected, and the commodity behaves like an asset, with no stockouts, whose price follows a random walk.

Perhaps the most notable feature of Table IV is the fact that the autocorrelation coefficients estimated for the harvest process are so close to the autocorrelation coefficients of the actual prices. Given that the storage model with i.i.d. harvests fails to generate the autocorrelation in

Table IV. Parameters for AR model with storage and actual autocorrelations of price

	a	b	δ	ρ	ρ actual
Cocoa	0·226	−0·064	0·0825	0·837	0·825
Coffee	0·245	−0·075	0·0355	0·833	0·797
Copper	0·535	−0·246	−0·0463	0·751	0·836
Cotton	0·862	−0·088	0·0104	0·944	0·901
Jute	0·526	−0·154	0·0534	0·741	0·728
Maize	0·731	−0·469	−0·0500	0·720	0·777
Palm oil	0·412	−0·244	−0·0500	0·864	0·740
Rice	0·609	−0·120	0·0273	0·889	0·852
Sugar	0·683	−0·493	0·1122	0·629	0·623
Tea	0·438	−0·053	0·2320	0·918	0·819
Tin	0·223	−0·038	0·0456	0·918	0·886
Wheat	0·680	−0·162	0·0508	0·785	0·878

commodity prices, an obvious topic of interest is whether a modestly autocorrelated harvest process could be supplemented by speculative storage so as to account for the autocorrelations in the data. But the results do not come out that way. Instead, the estimated coefficients are close to those in the data, and indeed for nine of the twelve commodities are actually estimated to be higher.

The left-hand panel of Figure 16 shows the actual price data for cotton, together with the one-period-ahead predictions from the full model and the AR(1). In this case, the two predictions are rarely sufficiently far apart for the difference to be perceptible, and this close degree of similarity between the two conditional expectations is common to all the commodities. Of course, the pseudo-likelihood depends not only on the conditional expectations but also on the conditional variances, which are plotted in the right-hand panel. The innovations

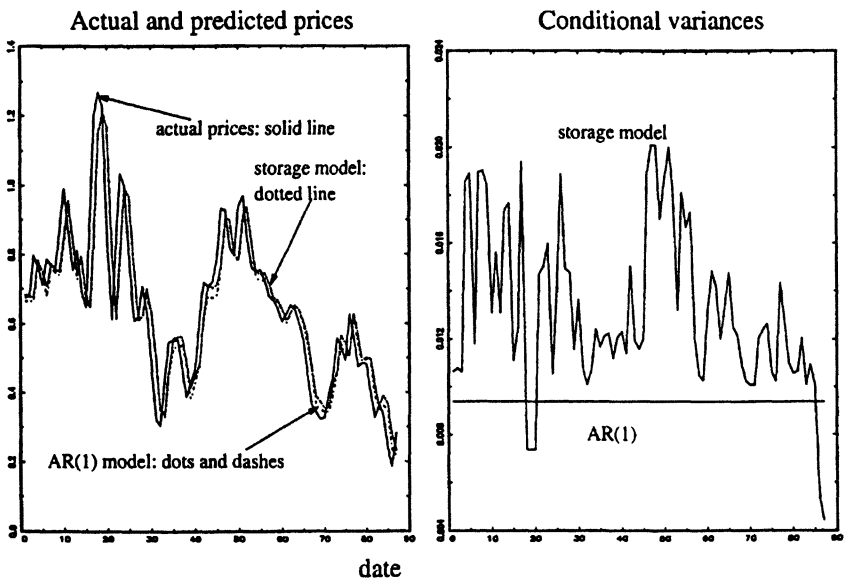

Figure 16. Cotton

of the AR process have constant variance by construction, while those for the full model show considerable fluctuation with the level of prices. The storage model is therefore capable of fitting better (or worse) than the AR(1), even when its step-ahead predictions are identical. If the conditional variance is high, a large prediction error will count less against the model than when the conditional variance is low, and some of the superiority of the full model over the AR(1) comes from this source.

As is further elaborated in the companion, substantive paper, we regard these results, although almost certainly subject to a wide margin of approximation error, to be sufficiently secure to give a preliminary judgment on the model. The i.i.d. model with storage seems like a very poor candidate for serious explanation of the data. The model with autoregressive shocks does better, but not a great deal better than a simple AR(1) without storage. However, the degree of autocorrelation implied by these models is too high to be consistent with one common interpretation, that the main source of shocks are agricultural supply shocks. Even for tree crops, where weather fluctuations affect harvests for several years, autocorrelations across years of 0·8 to 0·9 are quite implausible. In contrast, demand shocks, which are usually highly autocorrelated, are more plausibly consistent with our results. Under such an interpretation, storage seems to play only a small part in generating the autocorrelation in prices. From a methodological point of view, and even though allowing for storage improves the fit of a simple AR model, the result is still rather disappointing, since much of the complexity in the econometrics comes from handling the speculative storage.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under grant number SES-9223668. We are grateful to John Rust, Kenn Judd, and Mario Miranda for helpful comments on a previous draft.

REFERENCES

Berndt, E. K., B. H. Hall, R. E. Hall and J. A. Hausman (1974), 'Estimation and inference in nonlinear structural models', *Annals of Economic and Social Measurement*, **3**, 653–666.

Chambers, M. J. and R. E. Bailey, 'A theory of commodity price fluctuations', *Journal of Political Economy*.

Deaton, A. (1991), 'Saving and liquidity constraints', *Econometrica*, **59**, 1221–48.

Deaton, A. and G. Laroque (1992), 'On the behavior of commodity prices', *Review of Economic Studies*, **59**, 1–23.

Deaton, A. and G. Laroque (1994), 'Competitive storage and commodity price dynamics', Research Program in Development Studies, Princeton University and INSEE, Paris, processed (October).

Gallant, A. R. and G. Tauchen (1994), 'Which moments to match', Duke University, Raleigh, NC, processed.

Gourieroux, C., A. Monfort and A. Trognon (1984), 'Pseudo maximum likelihood methods: theory', *Econometrica*, **52**, 681–700.

Judd, K. L. (1992), 'Projection methods for solving aggregate growth model', *Journal of Economic Theory*, **58**, 410–452.

Laroque, G. and B. Salanie (1994), 'Estimating the canonical disequilibrium model: asymptotic theory and finite sample properties', *Journal of Econometrics*, **62**, 165–210.

Miranda, M. J. (1994), 'Maximum likelihood estimation of nonlinear rational expectations models by orthogonal polynomial projection methods', The Ohio State University, Columbus, OH, processed (December).

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.
Press, W. H., B. P. Flannery, S. E. Teukolsky and W. T. Vettering (1989), *Numerical Recipes: the art of scientific computing*, Cambridge University Press, Cambridge.
Rui, Xiongwen and M. J. Miranda (1995), 'Salvaging the rational expectations commodity storage model', The Ohio State University, Columbus, OH, processed (February).
Rust, J. (1987), Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher', *Econometrica*, **53**, 999–1033.
Rust, J. (1992), 'Do people behave according to Bellman's principle of optimality?' Hoover Institution, Stanford, CA, processed (May).
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.
Samuelson, P. A. (1971), 'Stochastic speculative price', *Proceedings of the National Academy of Sciences*, **68**, 335–337.
Schumaker, L. L. (1983), 'On shape preserving quadratic spline interpolation', *SIAM Journal on Numerical Analysis*, **20**, 854–64.
Tauchen, G. (1986), 'Finite state Markov chain approximations to univariate and vector autoregressions, *Economics Letters*, **20**, 177–81.
Tauchen, G. and R. Hussey (1991), 'Quadrature-based methods for obtaining approximate solutions to nonlinear asset-pricing models', *Econometrica*, **59**, 371–96.
Taylor, J. B. and H. Uhlig (1990), 'Solving nonlinear stochastic growth models: a comparison of alternative solution methods', *Journal of Business and Economic Statistics*, **8**, 1–17.
Williams, J. C. and B. D. Wright (1991), *Storage and Commodity Markets, Cambridge University Press, Cambridge.*