Discussion Paper 9

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
250 Marquette Avenue
Minneapolis, Minnesota 55480

January 1989

SOLVING A PARTICULAR GROWTH MODEL BY LINEAR QUADRATIC APPROXIMATION AND BY VALUE FUNCTION ITERATION

Lawrence J. Christiano

Federal Reserve Bank of Minneapolis

ABSTRACT

This paper studies the accuracy of two versions of the procedure proposed by Kydland and Prescott (1980, 1982) for approximating the optimal decision rules in problems in which the objective fails to be quadratic and the constraints linear. The analysis is carried out in the context of a particular example: a version of the Brock–Mirman (1972) model of optimal economic growth. Although the model is not linear quadratic, its solution can nevertheless be computed with arbitrary accuracy using a variant of the value function iteration procedures described in Bertsekas (1976). I find that the Kydland–Prescott approximate decision rules are very similar to those implied by value function iteration.

Prepared for presentation to the NBER Nonlinear Rational Expectations Modeling Study Group, Stanford, February 1988. I have benefited from the advice and suggestions of Terry Fitzgerald who also executed the computations.

1. Introduction

The linear–quadratic (LQ) method proposed by Kydland and Prescott (1980,1982) for approximating the solution to non–linear quadratic optimization problems has been applied by a wide variety of authors.[1] Little is known about the accuracy of this method in general. The purpose of this paper is to provide some evidence of its accuracy within the context of a particular example. This is done by comparing the LQ approximate solutions with the solutions obtained by discretizing the underlying state space and applying a variant of the value function iteration methods described in Bertsekas (1976). Since the grid for the endogenous variables in the state space is very fine, I expect that the solution obtained by value function iteration approximates very closely the solution in the version of the problem in which the endogenous state variables take on a continuum of values.

The example used in the paper is a version of the Brock Mirman (1972) one sector stochastic growth model. A solution to the model is a set of two decision rules. These relate end–of–period capital and current consumption to the current period state variables, respectively. There are two state variables, beginning–of–period capital and the current period's technology shock, which is a realization from a stationary stochastic process. Two versions of the linear–quadratic approximation are studied. The first relates end of period capital linearly to the state variables, and is called the linear LQ approximation. This is the original method used by Kydland and Prescott (1980,1982). The second approximation used is log linear in end–of–period capital and the state variables, and is therefore called the log–linear approximation method. It is applied in

[1]See, for example, Altug (1986), Christiano (1987c,1988), Cooley and Hansen (1988), Hansen (1985), Hansen and Sargent (1988) and King, Plosser and Rebelo (1988).

Christiano (1987a,c;1988) and King, Plosser and Rebelo (1988). This method has the virtue that in the special case in which the model of the paper does admit an analytic solution, then the log–linear decision rules and the exact decision rules coincide (see Proposition 1). This case is the model studied by Long and Plosser (1982) in which the depreciation rate on capital is 100%, the production function is Cobb–Douglas, and utility is logarithmic in consumption.

The accuracy of the approximate decision rules is evaluated on four dimensions. First, I compare the LQ decisions with those of the (approximately) exact solution at selected points in the state space. Then, I compare the LQ decision rules' implications for several first and second moments with those of the exact solution. This comparison is of particular relevance since first and second moments play an important role at the parameter selection and model evaluation stages, respectively, for many who use LQ approximations. Third, I report the amount, expressed as a fraction of the initial stock, that a planner who only knows the LQ decision rule, would be willing to pay to "learn" the exact decision rule. This is a measure of how close to optimal the LQ decision rules are. Finally, graphs of the steady state distribution of consumption and capital as implied by all three solutions are presented.

The results suggest that the LQ approximation is remarkably accurate for the example at hand. In addition, they show that the log–linear and linear LQ approximations are roughly equally accurate. This latter finding illustrates that the relative accuracy of the two decision rules is context specific. For example, in Christiano (1987a,b) and Christiano (1988,ftn.18) I show that in a model similar to the one in this paper, but with a productivity shock that is a logarithmic random walk, the two approximations are dramatically different. Christiano (1987a) shows that the difference reflects the accuracy of the log–linear approximation and the

very poor accuracy of the linear approximation in that context.

The plan of the paper is as follows. In section 2 I present the growth model that is studied, and its LQ approximate solution. Section 3 formulates the optimization problem as a dynamic programming problem and discusses its solution by value function iteration. Details of the algorithm appear in Appendix A. To my knowledge, the algorithm I use has not been used before. The appendix compares the computing times of the algorithm I use with other existing algorithms. Section 4 describes the model parameters used in the experiments. Section 5 presents the comparison of the LQ and value function iteration solutions. Section 6 concludes.

2. The Problem and Its Solution By LQ Approximation.

The problem I consider is the one good growth model in which the planner maximizes

$$(2.1) \qquad E_0 \Sigma_{t=0}^{\infty} (1-\tau)^{-1} C_t^{(1-\tau)} \beta^t$$

subject to

$$(2.2) \qquad C_t + K_t - (1-\delta)K_{t-1} = \exp(x_t)K_{t-1}^{\alpha}.$$

Here, C_t and K_t are date t consumption and the end–of–period t capital stock, respectively. Also, δ, α and τ are the rate of depreciation on capital, the share of income due to capital and risk aversion, respectively. We that assume that x_t is a realization from an n_x state, first order Markov chain with

$$(2.3a) \qquad x_t = \mathcal{X} = \{\mathcal{X}_1,...,\mathcal{X}_{n_x}\}$$
$$\text{Probability } \{x_{t+1} = \mathcal{X}_j | \; x_t = \mathcal{X}_i\} = \pi_{ij}$$
$$Ex_t = 0.$$

At times it is convenient to refer to the model for x_t in terms of its first order autoregressive representation:

$$(2.3b) \qquad x_t = \rho x_{t-1} + \epsilon_t,$$

where ϵ_t is mean zero with variance σ_ϵ^2 and is uncorrelated with x_{t-1}. Further

details about n_x, $\pi = [\pi_{ij}]$ and $\mathcal{S}$ are given below.

A solution to this problem is a function relating the date t decision, K_t, to the date t information variables, x_t and K_{t-1}. The exact solution is known only for the case $\tau = \delta = 1$ (see Long and Plosser [1982].) The rest of this section describes two variants of the linear quadratic (LQ) approximation method proposed by Kydland and Prescott (1980,1982). The first of these, the linear LQ method, approximates the decision rule for K_t by one that is linear in x_t and K_{t-1}. The second, the log–linear LQ method, approximates it by one in which the log of K_t is related linearly to x_t and the log of K_{t-1}. The decision rule delivered by the log–linear LQ method has the virtue that it coincides with the exact decision rule when $\delta = \tau = 1$.

2.a The Linear LQ Approximation.

Express the problem as a standard calculus of variations problem by substituting (2.2) into (2.1):

$$\text{(2.4)} \qquad \text{maximize } E_0 \, \Sigma_{t=0}^{\infty} \beta^t u(K_{t-1}, K_t, x_t)$$

subject to x_0 and K_{-1} given. Here,

$$\text{(2.5)} \qquad u(K,K',x) = \frac{1}{1-\tau} [\exp(x)K^{\alpha} + (1-\delta)K - K']^{(1-\tau)}.$$

The linear LQ method approximates the solution to (2.4) by the solution to the following linear quadratic optimization problem:

(2.6) maximize $E_0 \Sigma_{t=0}^{\infty} \beta^t U(K_{t-1}, K_t, x_t)$,

where U is the second order Taylor series expansion of u about $K_{t-1} = K_t = K^*$ and $x_t = x^*$. Here, K^* and x^* are the steady state values of K_t and x_t of the nonstochastic version of (2.4) obtained by setting $\epsilon_t = 0$ for all t. Trivially, $x^* = 0$. Also, it is easy to verify that

(2.7) $$K^* = \{\frac{\beta\alpha\exp(x^*)}{1-(1-\delta)\beta}\}^{[1/(1-\alpha)]}.$$

It is convenient to define $\tilde{K}_t = K_t - K^*$. At date t, the first order necessary condition for $\tilde{K}_t$ to solve (2.6) is

(2.8) $$E_t\tilde{K}_{t+1} - \phi\,\tilde{K}_t + \frac{1}{\beta}\tilde{K}_{t-1} = -(q/\beta)x_t.$$

Here,

(2.9) $$\phi = -\frac{u_{22} + \beta\, u_{11}}{\beta u_{12}} = 1 + \beta^{-1} + (\frac{1-\alpha}{\tau})[1 - (1-\delta)\beta](C^*/K^*),$$

where C^*/K^* is the steady state consumption to capital ratio, given by

(2.10) $$C^*/K^* = \frac{\beta^{-1} - 1 + \delta(1-\alpha)}{\alpha}.$$

In (2.9), u_{ij} is cross derivative of u with respect to its i^{th} and j^{th} arguments, evaluated at steady state. It can also be shown that

$$(2.11) \qquad q = \frac{u_{23} + \rho\beta u_{13}}{u_{12}} = \beta\{ (1-\rho)(\frac{C^*}{K^*} + \delta) + \frac{\rho\beta}{\tau}[\beta^{-1} - 1 + \delta]\frac{C^*}{K^*} \}K^*.$$

Let λ be the unique number which satisfies $|\lambda| \leq 1$ and $\lambda^2 - \phi\lambda + (1/\beta) = 0$. Then, the unique solution to (2.6) is the rule:

$$\tilde{K}_t = \lambda\tilde{K}_{t-1} + q\frac{\lambda}{1-\beta\rho\lambda}x_t,$$

or,

$$(2.12) \qquad K_t = (1-\lambda)K^* + \lambda K_{t-1} + q\frac{\lambda}{1-\beta\rho\lambda}x_t$$

$$= f_{linLQ}(K_{t-1},x_t),$$

say, where $f_{linLQ}(K,x) \equiv (1-\lambda)K^* + \lambda K + [q\lambda/(1-\beta\rho\lambda)]x$. Substituting this decision rule into (2.2), we get the linear LQ approximate decision rule for consumption, g_{linLQ}:

$$(2.13) \qquad C_t = g_{linLQ}(K_{t-1},x_t) \equiv \exp(x_t)K_{t-1}^{\alpha} + (1-\delta)K_{t-1} - f_{linLQ}(K_{t-1},x_t).$$

This decision rule gets its name from the fact that the decision rule for K_t is linear in its arguments. Clearly, g_{linLQ} is not itself linear.

2.b The Log–Linear LQ Approximation.

Let $k_t \equiv \log(K_t)$ and define:

(2.14) $\quad r(k,k',x) \equiv u(\exp(k),\exp(k'),x),$

where u is given in (2.5). Then, an equivalent way to write (2.4) is as follows:

(2.15) $\quad$ maximize $E_0 \Sigma_{t=0}^{\infty} \beta^t r(k_{t-1},k_t,x_t),$

with respect to decision rules for k_t. The Log–Linear LQ method approximates this decision rule by the one that solves

(2.16) $\quad$ maximize $E_0 \Sigma_{t=0}^{\infty} \beta^t R(k_{t-1},k_t,x_t),$

where R is the second order Taylor series expansion of r about $k_{t-1} = k_t = \log(K^*)$ and $x_t = x^*$. Let $\tilde{k}_t \equiv k_t - k^*$. The first order necessary condition for $\tilde{k}_t$ to solve (2.16) is

(2.17) $\quad E_t\tilde{k}_{t+1} - \phi\tilde{k}_t + (1/\beta)\tilde{k}_{t-1} = -[q/(\beta K^*)]x_t,$

where ϕ and q are as defined in (2.9) and (2.11), respectively.[2] The solution to (2.16), then, is

$$\tilde{k}_t = \lambda\tilde{k}_{t-1} + (q/K^*)\frac{\lambda}{1-\beta\rho\lambda}x_t,$$

[2]In this case, ϕ is $-(r_{22} + \beta r_{11})/(\beta r_{12})$, where r_{ij} is the cross derivative of r with respect to its i^{th} and j^{th} arguments, evaluated at $k_{t-1} = k_t = \log(K^*)$. The statement in the text follows from the facts $(r_{22} + \beta\ r_{11})/(\beta r_{12}) = (u_{22} + \beta\ u_{11})/(\beta u_{12})$ and $(r_{23} + \beta\rho r_{13})/r_{12} = (u_{23} + \beta\rho u_{13})/(u_{12}K^*)$.

or,

$$(2.18) \qquad K_t = (K^*)^{(1-\lambda)}\exp[\frac{q}{K^*}\frac{\lambda}{1-\beta\rho\lambda}x_t]\,K_{t-1}^{\lambda}$$

$$\equiv f_{loqLQ}(K_{t-1},x_t).$$

When $\tau = \delta = 1$, then $\lambda = \alpha$, $q/K^* = (1-\beta\rho\alpha)/\alpha$, and $(K^*)^{(1-\lambda)} = \alpha\beta$, so that (2.18) reduces to

$$(2.19) \qquad K_t = \alpha\beta\exp(x_t)K_{t-1}^{\alpha},$$

which is the exact solution to (2.4) (see Long and Plosser [1982].) This proves the following Proposition:

Proposition 1:

If $\tau = \delta = 1$,

then the Log–linear approximation is exact.

An analogous proposition can be proved for the linear LQ decision rule. In particular, when $\tau = \delta = 1$, then (2.12) reduces to $K_t = (1-\alpha)K^* + \alpha K_{t-1} + K^* x_t$. But, this is the first order Taylor series expansion of the right side of (2.19) about $x_t = 0$ and $K_{t-1} = K^*$. Thus, we have:

Proposition 2:

if $\tau = \delta = 1$,

then the linear LQ decision rule is the first order Taylor series expansion of the exact decision rule about the steady state values of x_t and K_{t-1}.

Denote the log–linear LQ decision rule by f_{logLQ}:

$$(2.20) \qquad K_t = f_{logLQ}(K_{t-1},x_t) = (K^*)^{(1-\lambda)}\exp\{x_t q\lambda/[K^*(1-\beta\rho\lambda)]\}\ K_{t-1}^{\lambda}.$$

The implied log–linear decision rule for consumption is g_{logLQ}:

$$(2.21) \qquad C_t = g_{logLQ}(K_{t-1},x_t) = \exp(x_t)K_{t-1}^{\alpha} + (1-\delta)K_{t-1} - f_{logLQ}(K_{t-1},x_t).$$

Unless $\tau = \delta = 1$, g_{logLQ} is not itself log linear.

3. The Solution by Value Function Iteration.

In the problem posed in the previous section only the exogenous shock, x_t, was assumed to lie on a discrete grid, $\mathscr{X}$. In particular, the capital stock was implicitly assumed to be able to take on a continuum of values. Value iteration methods require that the capital stock lie on a discrete grid and therefore do not, strictly speaking, apply to the problem posed in section 2. However, one expects that by choosing a sufficiently fine grid for K_t, $\mathscr{K}$, an arbitrarily accurate approximation to the underlying continuous problem can be obtained. In the calculations of this paper $\mathscr{K}$ is in fact extremely fine. The following discussion is informal. Details of the solution method used appear in the Appendix.

Problem (2.4) is expressed as a dynamic programming problem as follows:

$$(3.1) \qquad v(K_{t-1},x_t) = \max_{K_t \in A(K_{t-1},x_t)} \{u(K_{t-1},K_t,x_t) + \beta E_t v(K_t,x_{t+1})\},$$

subject to $\{x_t\}$ having Wold representation (2.3). Finite state Markov chains having this property will be described shortly. In (3.1), A denotes the feasible set of possible choices of K_t. Feasibility is determined by the non–negativity constraint on C_t and by the requirement that $K_t \in \mathscr{K}$. Formally:

$$(3.2) \qquad A(K,x) = \{K' \in \mathscr{K}: \exp(x)K^{\alpha} + (1-\delta)K - K' \geq 0\}.$$

The set $\mathscr{K}$ is a discrete interval of positive numbers which is described below.

One value function iteration method, *standard value iteration*, proceeds by starting with some initial v function on the right of (3.1), say v_j, then evaluating

the expression to the right of the equality in (3.1) and calling the result v_{j+1}. Then v_j is replaced by v_{j+1} and the procedure is repeated to yield v_{j+2}. This continues until there is little change in the v_j's. Call this converged v_j, v. In practice, it is possible to speed up the above contraction iterations along the lines suggested in the introduction (see the Appendix.) The function v is used to compute the decision rule for capital, f_{DP}, as follows:

$$(3.3) \qquad f_{DP}(K_{t-1},x_t) = \underset{K_t \in \Lambda(K_{t-1},x_t)}{\operatorname{argmax}} \{u(K_{t-1},K_t,x_t) + \beta E_t v(K_t,x_{t+1})\}.$$

The decision rule for C_t can then be obtained from f_{DP}, the decision rule for K_t, as follows:

$$(3.4) \qquad g_{DP}(K_{t-1},x_t) = \exp(x_t)K_{t-1}^{\alpha} + (1-\delta)K_{t-1} - f_{DP}(K_{t-1},x_t).$$

Let $\underline{x}$ and $\overline{x}$ denote the smallest and largest possible values of x_t, respectively. Also, let $\underline{K}$ be the limit of the sequence $\{K_i\}$, where K_0 is the smallest element in $\mathcal{K}$ and $K_i = f_{DP}(K_{i-1},\underline{x})$, i = 1,2,... . Similarly, $\overline{K}$ is the limit of the sequence K_j defined by the condition that K_0 is the largest element in $\mathcal{K}$ and $K_j = f_{DP}(K_{j-1},\overline{x})$, j = 1,2,... . Then, since $f_{DP}(\cdot,x)$ is increasing in x for all examples studied, it follows that $\mathcal{E}_{DP} \equiv \{K \in \mathcal{K}: \underline{K} \le K \le \overline{K}\}$ forms an ergodic set relative to the DP decision rule. That is, if $K_t \in \mathcal{K}$ should start out outside this set, it moves into it and once in it, it stays there. The steady state probability of K_t lying outside this set is zero. In a similar way, the linear and log–linear LQ decision rules

also imply ergodic sets, which I label $\mathscr{E}_{\text{linLQ}}$ and $\mathscr{E}_{\text{logLQ}}$, respectively.[3]

One way to choose the grid, $\mathscr{K}$, is to make its smallest (largest) point slightly less (greater) than the lowest (highest) value in $\mathscr{E}_{\text{DP}}$. Since knowledge of $\mathscr{E}_{\text{DP}}$ requires the decision rules, one could proceed by first getting a rough guess of $\mathscr{E}_{\text{DP}}$ based on decision rules obtained using a coarse grid, and based on $\mathscr{E}_{\text{linLQ}}$ or $\mathscr{E}_{\text{logLQ}}$. The second stage calculations can then be based on a very fine grid which contains few points outside $\mathscr{E}_{\text{DP}}$. This is the strategy followed for the calculations in this paper.

[3]Using (2.12), it is easy to confirm that $\mathscr{E}_{\text{linLQ}} = \{K^{*}+\underline{x}q\lambda/[(1-\beta\rho\lambda)(1-\lambda)]\ ,\ K^{*}+\bar{x}q\lambda/[(1-\beta\rho\lambda)(1-\lambda)]\}$. A similar calculation can be used to compute $\mathscr{E}_{\text{logLQ}}$ using (2.18).

4. Model Parameterization.

This section reports the model parameterizations used in the experiments. First, I describe the Markov Chain models used to model the exogenous shock.

4.a Markov Chain Models for $\{x_t\}$.

In the experiments described below, I used two Markov chain models for $\{x_t\}$: a two state model ($n_x = 2$) and a three state model ($n_x = 3$). The Markov chain is completely described by the state space of x_t, $\mathcal{X}$, and the transition probability matrix, π.

The Two State Markov Chain Model for $\{x_t\}$.

In this case

$$\pi = \begin{bmatrix} \phi & 1-\phi \\ 1-\phi & \phi \end{bmatrix}, \quad \mathcal{X} = \{-\sigma, \sigma\}. \tag{4.1}$$

The AR(1) representation associated with this Markov chain is (2.3) with

$$\rho = 2\phi - 1, \; E[\epsilon_t | x_{t-1}] = 0, \; E[\epsilon_t^2 | x_{t-1}] = \sigma_\epsilon^2 = \sigma^2(1-\rho^2). \tag{4.2}$$

Also, the steady state probabilities of $x_t = -\sigma$ and $x_t = \sigma$ are 1/2 each. Evidently, values for ρ and σ_ϵ^2 completely determine the parameters of the two-state Markov chain. In the experiments below, I set $\sigma_\epsilon^2 = (.01)^2$ and $(.10)^2$, and $\rho = .95$. This implies $\sigma = .032$ and .32, and $\phi = .975$. Prescott (1986,p.15) argues that a value of

σ_ϵ a little under .01 is empirically plausible. The large value of σ_ϵ (.10) was alsoused in order to get an idea of how large the shocks have to be before the LQ approximation deteriorates significantly. The large value of ρ corresponds well with Prescott's (1986) empirical finding that technology shocks are highly serially correlated.

The Three State Markov Chain Model for $\{x_t\}$

In this case

$$\pi = \begin{bmatrix} \phi & \gamma & 1-\phi-\gamma \\ \psi & 1-2\psi & \psi \\ 1-\phi-\gamma & \gamma & \phi \end{bmatrix}, \quad \mathcal{X} = \begin{bmatrix} -x \\ 0 \\ x \end{bmatrix}. \tag{4.3}$$

The Wold representation corresponding to this markov chain is also (2.3), and

$$\rho = 2\phi + \gamma - 1, \ E[\epsilon_t | x_{t-1}] = 0, \ \kappa = 1+.5\gamma/\psi, \ Var(x_t) = x^2/\kappa, \ \sigma_\epsilon^2 = Var(x_t)(1-\rho^2), \tag{4.4}$$

where $\kappa \equiv E(x_t^4)/[E(x_t^2)]^2$ is kurtosis.[4] The steady state probabilities for $x_t = -x$, $x_t = 0$, and $x_t = x$, are, respectively, P, (1–2P), and P, where $P = (2\kappa)^{-1}$. To determine this model, four parameters must be assigned values: ϕ, γ, ψ, and x. Thus, unlike in the two state case, values for σ_ϵ^2 and ρ are not sufficient. To determine values for this markov chain example, I set σ_ϵ^2, ρ, κ, and γ. In the

[4]Unlike in the two–state model, the innovation in the three state model is conditionally heteroscedastic.

experiments below, I set $\sigma_\epsilon^2 = (.01)^2$ and $(.10)^2$, $\rho = .95$, $\kappa = 3$, and $\gamma = .040$.[5] I

[5]In this case, for both values of σ_ϵ^2

$$\pi = \begin{bmatrix} .955 & .040 & .005 \\ .010 & .980 & .010 \\ .005 & .040 & .955 \end{bmatrix}.$$

set $\kappa = 3$ so that the model would resemble the normal distribution, for which $\kappa = 3$.

4.b Other Model Parameters.

I analyzed five parameterizations of the model. The first four set $\beta = .98$, $\tau = .5$, $\rho = .95$, $\alpha = .33$ and incorporate one of the four Markov chains models for x_t described above: the low ($\sigma_\epsilon = .01$) and high ($\sigma_\epsilon = .10$) variance two–state Markov chain and the low and high variance three–state Markov chain. Comparing results for these models permits judging their robustness to the number of states in the Markov chain and to the variance of the technology shock innovation. A fifth model was studied to determine the impact of risk aversion on the results. In that model, risk aversion is very high, with $\tau = 3$. Otherwise, the fifth model is parameterized in the same way as the model with $\tau = .5$ and a three state, low variance, Markov chain model for x_t.

The ergodic sets associated with each of the three capital decision rules corresponding to each of my five models are reported in Table 1. It is striking how similar $\mathcal{E}_{\text{logLQ}}$ and $\mathcal{E}_{\text{DP}}$ are. In contrast, $\mathcal{E}_{\text{linLQ}}$ is shifted to the left of $\mathcal{E}_{\text{DP}}$ in the high shock variance cases. Table 1 also reports the boundaries of the capital grid, $\mathcal{K}$, used when solving each model by value function iteration. In each case, $\mathcal{K}$ contains 20,000 points. The interval between grid points is reported in column 5 of Table 1. Column 6 of Table 1 reports the number of minutes of central processor time used in solving the model by value function iteration. The time used to solve the three state exogenous shock models exceeds by about 50 percent the time required for the two state exogenous shock models. This reflects that the number of points in the state space of the three shock models (60,000) exceeds that in the two

shock models (40,000) by 50 percent. In each case, the value function iterations were started with $v_0 \equiv 0$ and where considered to have converged when sup $\{|v_j - v_{j-1}|/|v_{j-1}|\}\times 100$ was less than .000001.

5. Comparison of LQ and DP Decision Rules.

This section reports comparisons of the LQ and DP decision rules, for each of the five models defined in section 4.b and column 1 of Table 1. There seems to be no best metric for comparing decision rules, and so I use several. The first compares the LQ and DP decision rules directly by tabulating their values at alternative points in the state space. The second compares several first and second moment properties of the decision rules. The third measures the amount an agent who uses an LQ decision rule would be willing to pay (expressed as a fraction of initial capital) to learn the DP decision rule. This is a direct measure of how close to optimal the LQ decision rules are. The fourth compares the steady state distribution of C_t and K_t implied by the LQ and DP decision rules.

5.a Tabulation of Decision Rules

Tables 2 and 3 tabulate the DP and two LQ approximate decision rules, at various points in the state space, and for the five models for which solutions were computed. The (K_{t-1},x_t) combinations represented in Tables 2 and 3 include all possible x_t's in the relevant markov chain and five representative K_{t-1}'s. Of these, the middle one is always K^* and the least and greatest ones are the end points of $\mathcal{E}_{DP}$, taken from Table 1. The other two points are half way between K^* and these end points. To aid in comparing the LQ and DP decision rules, cases where they differ by between 1 and 10 percent are indicated by a *, cases where they differ by between 10 and 20 percent are marked by a †, and cases where they differ by more than 20 percent are marked by a ††.

Consider first Table 2, which reports results for the two two state

Markov chain models. Panel A in that table shows that the DP rule and the two LQ decision rules are all approximately identical in the low variance case. In particular, if the capital and consumption decisions are rounded to one digit after the decimal the decisions are identical. Not surprisingly, the decision rules diverge somewhat for the high variance two state Markov model, results for which are reported in Panel B. The divergence is fairly minor in the case of the capital decision, where a difference exceeding 1 percent occurs only once. The differences are larger in the context of consumption, where the level of consumption implied by the LQ decision rules tend to overstate optimal consumption, in one case by over 20 percent.

It is also interesting to compare the decision rules according to whether they are increasing or decreasing in the state variables. In all cases in Table 2 the DP decision rules for C_t and K_t are increasing over the reported values of x_t and K_{t-1}. This property is shared by the LQ decision rules for K_t, a fact that can be verified analyticaly from the appropriate formulas in section 2.[6] Over the reported values of x_t, K_{t-1}, the LQ decision rules for C_t are also increasing in K_{t-1}. However, only the linear LQ decision rule for C_t is increasing in x_t. In particular, in the high variance case the log–linear decision rule for C_t is *decreasing* in x_t for $K_{t-1} = 86.14$ and 108.60.

Next consider Table 3, which reports results for the three state Markov chain models. Basically, the same picture that emerged from Table 2 emerges there as well. In particular, for the low variance version of the model with $\tau = .5$ (Panel

[6]The linear LQ approximate decision rule is not monotone in ϵ_t when $\rho = 1$ and the approximation is taken by first transforming the model so that the planner's choice variables are $c_t = C_t/\exp(x_t)$, $k_t = K_t/\exp(x_t)$. Then, even though k_t is monotone in ϵ_t, K_t is not. The log–linear LQ approximate decision rule for K_t, by contrast, is monotone in this case. For a fuller explanation and a demonstration of the quantitative significance of these differences, see Christiano (1987b;1988,ftn 18).

A), the DP and LQ decision rules are virtually identical. As in Table 2, the differences show up in the high variance case, principally in the consumption decision rule. Panel C shows that the high accuracy of the LQ decision rule when $\tau = .5$ and $\sigma_\epsilon = .01$ also obtains when $\tau = 3$.

All decision rules in Table 3 are monotone increasing over the reported values of the state variables, with the exception of the LQ decision rules for C_t in the high variance version of the model (Panel B.) Those rules are monotone increasing in K_{t-1}, but they fail to be monotone increasing in x_t. Since they differ in this respect from the corresponding DP rules, this reflects approximation error.

A feature of the three shock models that the two shock models lack is that both the low and high variance models with $\tau = .5$ share some common points in the state space. One of these, $(K_{t-1},x_t) = (63.69,0.0)$, is reported in Table 3. Because their construction imposes certainty equivalence, the LQ decisions for C_t and K_t at this point is the same for both the low and high variance models. However, the exact problem does not satisfy certainty equivalence, and so there is no reason to expect the DP rules to share this property. In fact, Table 3 indicates that $f_{DP}(63.69,0.0)$ is 63.69 and 63.68 for the low and high shock models, respectively. Also, $g_{DP}(63.69,0.0)$ is 3.94 and 3.95 in these two cases. Thus, while certainty equivalence does not to hold exactly, it appears to do so approximately.[7] This may be one of the reasons why the LQ approximations are so accurate.

In sum, the evidence in Tables 2 and 3 suggest that for reasonable shock variance (eg., the low variance case), the LQ approximation is very accurate, even with high risk aversion, $\tau = 3$. When the shock variances get very large, then—not

[7]A factor which complicates interpretation of this exercise is that although it involves a mean–preserving spread on x_t, the experiment does not involve a mean preserving spread on $\exp(x_t)$. By inducing a mean–preserving spread in x_t, the result describes the response to the shift in the mean of the distribution of $\exp(x_t)$, in addition to its variance.

surprisingly—the quality of the approximations begin to deteriorate. Based only on the evidence in Tables 2 and 3, it is hard to say which approximation—the linear, or log–linear LQ—performs better in the high variance case. On the one hand, there is evidence that the log–linear LQ approximation performs poorly at more points in the state space. For example, there are more †'s associated with the logLQ results than with the linLQ results in Tables 2 and 3. On the other hand, there is evidence that the states in which the log–linear LQ decision rules perform worst have lower probability than the states in which the linear LQ decision rules perform poorly. This is suggested by the fact that the logLQ decision rules perform worst in states with (low K, high x) and (high K, low x) combinations, whereas the reverse is true for the linLQ decision rules. Given that K_{t-1} and x_t are positively correlated (which the must be, given the high positive autocorrelation of x_t), then—other things equal—this would cause the log–linear decision rule to dominate the linear one in a weighted overall sense.

5.b First and Second Moment Implications of LQ and DP Decision Rules

Tables 4 – 7 report selected first and second moment properties of the DP and LQ decision rules, obtained by Monte Carlo simulation. I simulated 100 data sets on C_t, K_t, $Y_t \equiv C_t + K_t - K_{t-1}$, the risk free rate of interest, R_t, and the marginal product of capital, $MP_{k,t}$. Each data set has length 10,050, but the first 50 observations were discarded prior to computing first and second moments. The risk free rate of interest, R_t, is defined in the usual way, as $R_t = u'(C_t)/[\beta E_t u'(C_{t+1})] - 1$, where the conditional expectation is evaluated relative to the appropriate consumption decision rule and Markov chain probabilities, and $u(C_t) \equiv C_t^{(1-\tau)}/(1-\tau)$. The marginal product of capital is $MP_{k,t} \equiv \theta Y_{t+1}/K_t$,

where $Y_t \equiv \exp(x_t)K_{t-1}^{\alpha}$ is output.

In executing the simulations I found that the linear LQ decision rule occasionally implies a negative value for C_t. This happened only when the exogenous shocks were drawn from the three–state, high variance markov chain. To accommodate this, I redefined f_{linLQ} and g_{linLQ} in such a way that whenever they implied a negative C_t, C_t was set to .01 and K_t was adjusted appropriately. These redefined linear LQ decision rules where also used in computing the risk free rate of interest. Of the 1 million total values of C_t computed, 92 had to be adjusted in this way. Similarly, negative values of C_t where encountered in computing .67% of the R_t's.

First Moments

Table 4 reports first moment properties of the DP and LQ decision rules for the $\tau = .5$ models, as indicated in the column headings. The first column contains the variable whose mean is reported in the remaining columns. Those columns contain the average, across 100 simulations, of the mean value of the variable. Numbers in parentheses in columns 2 – 13 are the standard deviation across the 100 simulations. The small size of the standard deviations reflects the large number of observations per simulation (10,000.) Numbers in parentheses in the first column are steady state values of the associated variable. Comparison of the first moment properties of the DP rules with the corresponding steady states permits assessing an assumption implicit in many applications that utilize the LQ approximation. This assumption—that steady states and unconditional means roughly coincide—plays a role in two places in applied work. First, there would be little sense in approximating a model about steady state if the model's variables did

not fluctuate about this point in the stochastic version of the problem. Second, many empirical researchers who use the LQ approximation select model parameter values by equating nonstochastic steady state properties of their model with corresponding sample statistics (Christiano [1988] and Kydland and Prescott [1982].) This method of assigning parameter values would be inappropriate if the nonstochastic steady state diverged substantially from the mean of the stochastic version of the model.

Consider first capital, consumption and output. Table 4 indicates that, for the high variance economy, the mean of these variables is roughly 10% higher than their steady state values. In addition, the mean of the capital output ratio is about 7% higher than its steady state value. Presumably, the larger average capital stock in the stochastic economy reflects households' efforts to insure themselves against the risk associated with the production technology. Recall, however, that the innovation to the technology shock in the high variance economy is more than 10 times what is plausible empirically. In the empirically more plausible low variance economy, $EK_t/Y_t \cong K^*/Y^*$, $EK_t \cong K^*$, $EC_t \cong C^*$ and $EY_t \cong Y^*$ (stars indicate steady state quantities.)

The mean value of capital implied by the linear LQ decision rule is roughly equal to K^*, as it must given that it is linear. The log–linear LQ decision rule implies a larger mean value of K_t in the high variance economies because of the convexity of the exponential function. Thus, in the high variance economies, the mean of the logLQ capital stock lies in between that of the DP and linLQ decision rules. The same is true for C, Y and K/Y.

Next consider R_t and $MP_{k,t}$. In nonstochastic steady state these quantities are both $\beta^{-1} - 1 = .0204$. In the stochastic version of the model one expects $EMP_{k,t} > ER_t$. This reflects that $MP_{k,t}$ is the return on a riskier

investment than is R_t since the states in which the former pays off the most are those in which consumption is valued least, ie., $Cov_t(MP_{k,t}, u'(C_{t+1})) < 0$. As it turns out, both ER_t and $EMP_{k,t}$ are approximately $\beta^{-1} - 1$, even in the high variance model. The fact that the average equity premium, $E[MP_{k,t} - R_t]$, is roughly zero in this model is reminiscent of a similar result obtained by Mehra and Prescott (1985) for an endowment economy.

In sum, the evidence in Table 4 for the four $\tau = .5$ models suggests that the steady state properties of the nonstochastic version of the model approximate closely the corresponding first moment properties of the stochastic version of the model, as long as the innovation variance to the technology shock is of plausible magnitude (i.e., $\sigma_\epsilon = .01$.) Table 7 reports first moment results for the $\tau = 3$, low variance model. Like in the $\tau = .5$, low variance models, there is little difference between steady states and unconditional means in the $\tau = 3$ model.

Second Moment Properties

Tables 5 – 7 report second moment properties of the models. There, σ_w denotes the standard deviation of the variable w_t. In addition, $\rho_{x,y}(\tau)$ denotes the correlation between x_t and $y_{t-\tau}$, and Δc_t signifies $C_t - C_{t-1}$. Numbers not in parentheses are the average of the associated statistic, across the 100 data sets. Numbers in parentheses are the corresponding standard deviation.

Consider first Table 5, which reports results for the $\tau = .5$ models. A striking feature of that Table is that the results differ so little between decision rules and models. The only quantitatively meaningful differences lie in σ_r/σ_y and $\rho_{r,\Delta c}(0)$. With regard to the former the standard deviation of σ_r/σ_y is higher for the linLQ decision rule and the high variance, three shock economy than for the

other cases. With regard to $\rho_{r,\Delta c}(0)$, the LQ versions of that statistic are smaller in the high variance models while the DP versions appear scale independent.[8] Since the LQ and DP versions are equal for the low variance economies and scale independence seems plausible, I interpret this to reflect approximation error in the LQ approximation.

Next, consider the dynamic correlations for the $\tau = .5$ economies reported in Table 6. As in the other tables, there are few significant discrepancies between solution methods and models. One discrepancy is that the correlations based on the linLQ solution to the high variance, three state markov model are all smaller than the other correlations, presumably reflecting approximation error. Another discrepancy is the scale dependence in the LQ versions of that $\rho_{\Delta c,\Delta c}(i)$, i=1,2.

Table 7 contains the second moment results for the $\tau = 3$ model. There are virtually no noticeable discrepancies between LQ and DP second moments. One exception is $\rho_{\Delta c,r}(2)$, which is lower for the DP decision rule than for the LQ decision rules.

An interesting feature of the results in Table 7 is that the correlations between consumption changes and lagged variables is close to zero when $\tau = 3$ and much further from zero in the $\tau = .5$ case. In this respect, the $\tau = 3$ results are close to what I found in a version of this model that I have studied elsewhere (see Christiano [1987a,c]) in which $\rho = \tau = 1$, $\delta = .018$, $\beta = .99$, and in which hours are variable. In that model, consumption changes are also approximately uncorrelated with lagged variables.

---

[8]The importance of the very fine grid used in the paper showed up in these second moment calculations. For example, when I computed the DP decision rules with a grid of .01 between capital points, then I got the following results for $\rho_{r,\Delta c}(0)$ in the four $\tau = .5$ models: .787 (.006), .365 (.006), .780 (.007) and .406 (.013) in the high and low variance, two state models and the high and low variance three state models, respectively. Thus, using a grid coarser than the one underlying the results in Table 5 I found that the DP rules imply some scale dependence.

To summarize, in the low variance economies the first and second moment properties implied by the LQ approximations and the value function iteration solution are roughly identical, even with high risk aversion. Discrepancies occur for very large shock variances. An interesting feature of the results is that second moment properties seem relatively insensitive to whether the exogenous shock is drawn from a two or three state Markov chain.

5.c The Value of the DP Rule to an LQ Decision Maker

Table 8 reports the amount, as a fraction of K_{t-1}, that a planner using an LQ decision rule would be willing to pay to learn the DP rule in the four $\tau = .5$ models. This quantity was computed for the same (K_{t-1},x_t) combinations used in Tables 2 and 3. Following is a discussion of how this was done.

In order to place the LQ and DP decision rules on a comparable basis, I redefined the LQ rules slightly. Instead of allowing f_{logLQ} and f_{linLQ} to map (K_{t-1},x_t) onto the real line, for purposes of the computations in Table 8, $f_{logLQ}(K_{t-1},x_t)$ and $f_{linLQ}(K_{t-1},x_t)$ were replaced by the nearest point in $A(K_{t-1},x_t)$, defined in (3.2). Given the fine grid, $\mathcal{K}$, this adjustment presumably has negligible effect.

I computed the v_{logLQ} and v_{linLQ} functions that solve the following functional equations:

$$(5.1) \qquad v_{logLQ}(K_{t-1},x_t) = u(K_{t-1},f_{logLQ}(K_{t-1},x_t),x_t) + \beta E_t v_{logLQ}(K_t,x_{t+1})$$

and

$$(5.2) \qquad v_{linLQ}(K_{t-1},x_t) = u(K_{t-1},f_{linLQ}(K_{t-1},x_t),x_t) + \beta E_t v_{linLQ}(K_t,x_{t+1}),$$

for u defined in (2.5). In each case, the expectation operator was evaluated relative to the relevant 2– or 3– state Markov chain with high or low variance (I do not index the v functions by the Markov chain model in order to avoid complicating the notation.)

Relative to a given model (defined by the markov chain for the exogenous shock) and specified initial conditions, I computed the loss of using the LQ decision rules as follows. Let $v^* = v_{logLQ}(K_{t-1},x_t)$. Then, let K' be defined by the property $v_{DP}(K',x_t) = v^*$, ie., K' is that level of capital ($K' \leq K_{t-1}$) such that a planner starting with K', x_t and knowing the DP rule achieves the same utility as a planner starting with K_{t-1}, x_t who uses the log–LQ decision rule. Evidently, the LQ planner would be willing to pay no more than $K_{t-1}-K'$ in order to acquire knowledge of the DP rule. The table reports this as a percent of K_{t-1}, ie., $[(K_{t-1}-K')/K_{t-1}]\times 100$. Similar calculations were carried out for the linear LQ decision rule.

It is not surprising, in view of the preceding results, that for the low variance shock distributions with $\tau = .5$, the LQ planner would not pay anything to acquire the DP decision rule (see Table 8, panels A and C.) Thus, for these shock distributions, the LQ rules are roughly optimal. I obtained exactly the same results for the low variance, $\tau = 3$ model as for the low variance, $\tau = .5$ models.

With high shock variances, the results in Panels B and D show that the LQ planner would pay a positive, though still very small, amount to acquire the DP rule. Except when very far from steady state, the planner would pay less than 1% of initial capital.

5.d Steady State Capital and Consumption Distribution

Figures 1a and b plot steady state capital and consumption, respectively, for the high variance, two state markov chain model for $\tau = .5$. The same is done in Figures 2a and b for the high variance, three state markov chain model. The low variance steady state distributions are not plotted because they actually coincide.

A distinguishing feature of these figures is the bi–modal distribution when the Markov chain has two states and the uni–modal distribution for the three state markov process. In the two state case, note the pronounced lack of symmetry in all but one of the distributions. This reflects the nonlinearity of all but one of the decision rules. The exception, f_{linLQ}, produces a roughly symmetric steady state distribution for capital.

An interesting feature of these charts is the left shift in the linLQ distributions for C_t and K_t relative to the logLQ and DP distributions. In view of this, it is not surprising that the non–negativity constraint on C_t proved to be occasionally binding when the three state, high variance Markov chain version of the model was solved by linear LQ approximation.

6. Conclusion.

The purpose of this paper was to evaluate the accuracy of two methods for approximating the consumption and capital decision rules that solve a version of the Brock and Mirman (1972) optimal growth problem. The strategy taken was to define the growth problem in such a way that numerical dynamic programming methods could be used to obtain arbitrarily accurate approximations to the exact decision rules. This involved, first, positing a discrete distribution for the exogenous shocks of the model and, second, forcing the capital stock to lie on a very fine grid. The accuracy of the LQ approximate decision rules was evaluated by comparing them along four dimensions with the presumed exact solutions obtained by dynamic programming (DP) methods.

I found that the LQ approximation works well in the model economy studied in this paper. In this respect the conclusions are similar to those reached by others (eg., Christiano [1986,1987a] and Danthine, Donaldson and Mehra [1988]), who looked at different examples.

An interesting feature of the model economies considered is that increased risk aversion widens the ergodic set for capital and reduces the correlation between consumption changes and lagged variables. It would seem worthwhile to explore the economics underlying this result.

Table 1: Decision Rule Information[1]

Model (1)	$\mathcal{E}_{\text{logLQ}}$ (2)	$\mathcal{E}_{\text{linLQ}}$ (3)	$\mathcal{E}_{\text{DP}}$ (4)	$\mathcal{K}$ (5)	Increment (6)	CPU Time (7)
Two-state, $\sigma_\epsilon = .10$ $\tau = .5$	{37,110}	{29,98}	{37,109}	{35,115}	.00400	213.84
Two-state, $\sigma_\epsilon = .01$ $\tau = .5$	{60,67}	{60,67}	{60,67}	{55,70}	.00075	214.02
Three-state, $\sigma_\epsilon = .10$ $\tau = .5$	{25,163}	{4,123}	{25,161}	{20,165}	.00725	310.58
Three-state, $\sigma_\epsilon = .01$ $\tau = .5$	{58,70}	{58,70}	{58,70}	{55,75}	.00100	314.80
Three-state, $\sigma_\epsilon = .01$ $\tau = 3.0$	{49,83}	{47,80}	{49,83}	{45,85}	.00200	342.08

[1]Column 1: Number of states and σ_ϵ value indicate the Markov chain model of the x_t's. The only other parameter that differs between models is τ.
Columns 2–4: First and second numbers associated with $\mathcal{E}$ are the upper and lower boundaries, respectively, of the ergodic set associated with the linLQ, logLQ, or DP capital decision rules, as indicated by the subscript.
Column 5: The first and second numbers associated with $\mathcal{K}$ are the boundaries of the capital grid used in the value function iteration calculations.
Column 6: Increment between adjacent values of capital in $\mathcal{K}$.
Column 7: Time, in central processing unit minutes, used to solve the associated model by value function iteration on the Federal Reserve Bank of Minneapolis' Amdahl dual 580 mainframe computer. Details of the solution method are described in the Appendix. In the case of the first four models, p = 10 for all j, using notation presented in the appendix. In the case of the last model, p = 10 for j = 1,...,108 and p = ∞ for j = 109,...,118, whereupon convergence occured. The convergence criterion is reported in section 4b.

Table 2: Capital and Consumption Rules Two State Markov Process[1] [2]

Panel A: $\sigma_\epsilon = .01,\ \tau = .5$

K_{t-1} grid ↓	x_t=-.03	x_t=.03	x_t=-.03	x_t=.03	x_t=-.03	x_t=.03
	f_{DP}		f_{logLQ}		f_{linLQ}	
60.32	60.32	60.53	60.32	60.52	60.31	60.53
62.00	61.95	62.16	61.95	62.16	61.95	62.16
63.69	63.58	63.79	63.58	63.79	63.58	63.79
65.46	65.30	65.51	65.29	65.51	65.30	65.51
67.23	67.01	67.23	67.01	67.23	67.02	67.23
	g_{DP}		g_{logLQ}		g_{linLQ}	
60.32	3.75	3.78	3.75	3.79	3.75	3.78
62.00	3.83	3.87	3.83	3.87	3.84	3.87
63.69	3.92	3.96	3.92	3.96	3.92	3.96
65.46	4.01	4.05	4.02	4.05	4.01	4.05
67.23	4.10	4.14	4.11	4.14	4.10	4.14

Panel B: $\sigma_\epsilon = .10,\ \tau = .5$

K_{t-1} grid ↓	x_t=-.32	x_t=.32	x_t=-.32	x_t=.32	x_t=-.32	x_t=.32
	f_{DP}		f_{logLQ}		f_{linLQ}	
36.78	36.78	38.63	36.79	38.05*	36.55	38.70
50.24	49.79	51.82	49.76	51.47	49.58	51.73
63.69	62.77	64.95	62.62	64.77	62.61	64.76
86.19	84.45	86.85	83.95	86.83	84.41	86.56
108.69	106.11	108.69	105.11	108.72	106.21	108.36
	g_{DP}		g_{logLQ}		g_{linLQ}	
36.78	2.39	2.68	2.38	3.25††	2.62*	2.61*
50.24	3.09	3.43	3.12	3.78†	3.30*	3.52*
63.69	3.78	4.16	3.93*	4.34*	3.93*	4.35*
86.19	4.90	5.33	5.40†	5.35	4.94	5.62*
108.69	5.99	6.47	6.99†	6.45	5.89*	6.80*

[1]The table reports capital and consumption decisions for various points in the state space based on the DP, logLQ and linLQ decision rules. Rows correspond to values of initial capital and columns correspond to technology shock values, as indicated.

[2]Let z denote the ratio of an LQ decision to a DP decision at a given point in the state space. Let z' be $|(z-1)\times 100|$, rounded to the nearest integer, where $|\cdot|$ denotes the absolute value operator. Then, '*' indicates $1 \leq z' < 10$, '†' indicates $10 \leq z' < 20$, and '††' indicates $z' > 20$.

Table 3: Capital and Consumption Rules Three State Markov Process[1]

Panel A: $\sigma_\epsilon = .01$, $\tau = .5$

K_{t-1} ↓	x_t=−.06	x_t=.0	x_t=.06	x_t=−.06	x_t=.0	x_t=.06	x_t=−.06	x_t=.0	x_t=.06
	f_{DP}			f_{logLQ}			f_{linLQ}		
57.96	57.96	58.14	58.32	57.96	58.13	58.30	57.95	58.14	58.32
60.82	60.73	60.91	61.10	60.73	60.91	61.09	60.73	60.91	61.10
63.69	63.51	63.69	63.88	63.50	63.69	63.87	63.50	63.69	63.87
66.82	66.54	66.72	66.92	66.53	66.72	66.92	66.54	66.73	66.91
69.96	69.58	69.76	69.96	69.55	69.76	69.96	69.58	69.77	69.95
	g_{DP}			g_{logLQ}			g_{linLQ}		
57.96	3.61	3.64	3.67	3.61	3.65	3.70	3.62	3.64	3.67
60.82	3.76	3.79	3.82	3.76	3.79	3.83	3.77	3.79	3.82
63.69	3.91	3.94	3.97	3.91	3.94	3.98	3.91	3.94	3.98
66.82	4.07	4.10	4.13	4.08	4.10	4.13	4.07	4.10	4.14
69.96	4.23	4.26	4.29	4.25	4.27	4.29	4.23	4.26	4.30

Panel B: $\sigma_\epsilon = .10$, $\tau = .5$

K_{t-1} ↓	x_t=−.55	x_t=.0	x_t=.55	x_t=−.55	x_t=.0	x_t=.55	x_t=−.55	x_t=.0	x_t=.55
	f_{DP}			f_{logLQ}			f_{linLQ}		
24.60	24.60	25.65	27.53	24.61	25.34*	26.09*	23.95*	25.81	27.68
44.14	43.45	44.71	46.97	43.36	44.65	45.97*	42.89*	44.75	46.61
63.69	62.27	63.68	66.20	61.85	63.69	65.58	61.82	63.69	65.55
112.26	108.95	110.63	113.64	107.12*	110.30	113.57	108.89	110.75	112.61
160.84	155.60	157.48	160.84	151.76*	156.26	160.90	155.95	157.81	159.67
	g_{DP}			g_{logLQ}			g_{linLQ}		
24.60	1.65	1.83	2.08	1.64	2.14†	3.52††	2.30††	1.66*	1.93*
44.14	2.69	2.92	3.25	2.78*	2.98*	4.25††	3.26††	2.88*	3.61†
63.69	3.68	3.95	4.34	4.10†	3.94	4.97†	4.12†	3.94	5.00†
112.26	6.04	6.38	6.90	7.87††	6.71*	6.96	6.10*	6.26*	7.92†
160.84	8.30	8.70	9.31	12.15††	9.92†	9.25	7.96*	8.37*	10.47†

Panel C: $\sigma_\epsilon = .01$, $\tau = 3.0$

K_{t-1} ↓	x_t=−.06	x_t=.0	x_t=.06	x_t=−.06	x_t=.0	x_t=.06	x_t=−.06	x_t=.0	x_t=.06
	f_{DP}			f_{logLQ}			f_{linLQ}		
48.95	48.95	49.09	49.23	48.95	49.07	49.19	48.93	49.09	49.24
56.32	56.25	56.39	56.54	56.25	56.38	56.52	56.23	56.39	56.54
63.69	63.54	63.69	63.85	63.53	63.69	63.84	63.53	63.69	63.84
73.24	72.99	73.15	73.32	72.96	73.14	73.32	72.99	73.15	73.30
82.78	82.44	82.61	82.78	82.38	82.58	82.78	82.45	82.61	82.76
	g_{DP}			g_{logLQ}			g_{linLQ}		
48.95	3.42	3.48	3.54	3.42	3.49	3.58	3.43	3.47	3.53
56.32	3.65	3.71	3.78	3.65	3.72	3.80	3.66	3.71	3.78
63.69	3.87	3.94	4.00	3.88	3.94	4.01	3.88	3.94	4.01
73.24	4.15	4.21	4.28	4.17	4.22	4.28	4.14	4.21	4.29
82.78	4.41	4.47	4.54	4.46	4.49	4.54	4.39	4.47	4.56

[1]See notes to table 2.

Table 4

First Moment Properties: $\tau = .5$ Economy*

	Two-State $\sigma_\varepsilon = .10$			Two-State $\sigma_\varepsilon = .01$			Three-State $\sigma_\varepsilon = .10$			Three-State $\sigma_\varepsilon = .01$		
Statistic**	DP	Log	Lin	DP	Log	Lin	DP	Log	Lin	DP	Log	Lin
C (3.94)	4.34 (.15)	4.27 (.15)	4.19 (.15)	3.94 (.014)	3.94 (.014)	3.94 (.014)	4.31 (.13)	4.24 (.13)	4.16 (.13)	3.94 (.012)	3.94 (.012)	3.94 (.012)
Y (3.94)	4.34 (.15)	4.27 (.15)	4.19 (.15)	3.94 (.014)	3.94 (.014)	3.94 (.014)	4.31 (.13)	4.24 (.13)	4.16 (.13)	3.94 (.012)	3.94 (.012)	3.94 (.012)
K (63.69)	70.85 (2.64)	67.61 (2.65)	64.00 (2.53)	63.78 (.25)	63.75 (.25)	63.72 (.25)	70.44 (2.34)	67.06 (2.25)	63.42 (2.10)	63.73 (.21)	63.70 (.21)	63.66 (.21)
K/Y (16.17)	17.30 (.068)	16.73 (.059)	16.08 (.103)	16.18 (.005)	16.18 (.005)	16.17 (.005)	17.29 (.098)	16.71 (.006)	15.99 (.131)	16.18 (.004)	16.17 (.004)	16.17 (.005)
MP_k (.0204)	.0204 (.65E-4)	.0211 (.78E-4)	.0220 (.13E-3)	.0204 (.66E-5)	.0204 (.67E-5)	.0204 (.67E-5)	.0204 (.55E-4)	.0211 (.86E-4)	.0224 (.35E-3)	.0204 (.56E-5)	.0204 (.57E-5)	.0204 (.60E-5)
R (.0204)	.0204 (.88E-4)	.0203 (.12E-3)	.0204 (.94E-4)	.0204 (.90E-5)	.0204 (.89E-5)	.0204 (.90E-5)	.0204 (.74E-4)	.0203 (.14E-3)	.0200 (.72E-3)	.0204 (.75E-5)	.0204 (.75E-5)	.0204 (.75E-5)

*Results based on 100 data sets each of length 10,000, using the indicated solution (DP, log-linear LQ, or linear LQ) to the version of the growth model which incorporates the indicated probability model for the exogenous shock, x_t (two- or three-state Markov chain, with high or low variances). Initial conditions were randomized across simulations. Standard deviations across simulations appear in parentheses. For parameter values, see section 3 in the text.

**C,Y,K,R,MP_k denote the means of C_t, Y_t, K_t, R_t, and $MP_{k,t}$, respectively. Numbers in parentheses in this column are steady state values.

Table 5

Second Moments: $\tau = .5$ Economy*

	Two-State $\sigma_\varepsilon = .10$			Two-State $\sigma_\varepsilon = .01$			Three-State $\sigma_\varepsilon = .10$			Three-State $\sigma_\varepsilon = .01$		
Statistic	DP	Log	Lin	DP	Log	Lin	DP	Log	Lin	DP	Log	Lin
σ_c/σ_y	.78 (.02)	.79 (.02)	.78 (.02)	.78 (.02)	.78 (.02)	.78 (.02)	.77 (.02)	.78 (.02)	.80 (.02)	.78 (.02)	.78 (.02)	.78 (.02)
σ_{dk}/σ_y	.56 (.02)	.57 (.02)	.54 (.02)	.55 (.02)	.55 (.02)	.55 (.02)	.56 (.02)	.54 (.02)	.51 (.02)	.55 (.02)	.55 (.02)	.55 (.02)
σ_r/σ_y	.003 (.0001)	.003 (.0001)	.004 (.0002)	.003 (.0001)	.003 (.0001)	.003 (.0001)	.003 (.0001)	.003 (.0002)	.002 (.017)	.003 (.0001)	.003 (.0001)	.003 (.0001)
σ_y	1.62 (.02)	1.61 (.02)	1.59 (.02)	.155 (.002)	.155 (.002)	.155 (.002)	1.77 (.10)	1.74 (.10)	1.69 (.08)	.155 (.007)	.155 (.007)	.155 (.007)
$\rho_{r,c}(0)$	-.05 (.01)	-.08 (.03)	-.06 (.01)	-.05 (.013)	-.05 (.013)	-.05 (.013)	-.05 (.02)	-.12 (.03)	-.05 (.07)	-.05 (.016)	-.05 (.016)	-.05 (.017)
$\rho_{dk,c}(0)$	.11 (.005)	.06 (.010)	.11 (.007)	.10 (.005)	.10 (.005)	.10 (.005)	.11 (.008)	.09 (.015)	.13 (.022)	.10 (.007)	.10 (.007)	.10 (.007)
$\rho_{r,\Delta c}(0)$	.80 (.006)	.49 (.015)	.55 (.014)	.81 (.004)	.82 (.007)	.82 (.005)	.78 (.007)	.32 (.019)	.01 (.073)	.81 (.005)	.81 (.007)	.80 (.008)

*σ_x denotes the standard deviation of $\{X_t\}$. $\rho_{x,y}(\tau)$ denotes the correlation between X_t and $Y_{t-\tau}$. Δc_t denotes $C_t - C_{t-1}$, $dk_t \equiv K_t - K_{t-1}$. See also the notes to Table 4.

Table 6 Dynamic Correlations: $\tau = .5$ Economy*

	Two-State $\sigma_\varepsilon = .10$			Two-State $\sigma_\varepsilon = .01$			Three-State $\sigma_\varepsilon = .10$			Three-State $\sigma_\varepsilon = .01$		
Statistic	DP	Log	Lin	DP	Log	Lin	DP	Log	Lin	DP	Log	Lin
$\rho_{\Delta c,r}(1)$	.53 (.014)	.51 (.027)	.47 (.022)	.54 (.014)	.54 (.016)	.54 (.014)	.52 (.016)	.39 (.021)	.14 (.065)	.54 (.018)	.54 (.018)	.53 (.020)
$\rho_{\Delta c,r}(2)$	.48 (.016)	.45 (.024)	.43 (.020)	.49 (.017)	.50 (.018)	.50 (.017)	.48 (.018)	.34 (.020)	.13 (.057)	.49 (.019)	.49 (.020)	.49 (.021)
$\rho_{\Delta c,\Delta c}(1)$	.44 (.014)	.23 (.021)	.30 (.022)	.44 (.014)	.44 (.017)	.45 (.014)	.45 (.018)	.13 (.022)	.08 (.013)	.44 (.018)	.44 (.019)	.43 (.021)
$\rho_{\Delta c,\Delta c}(2)$	.40 (.016)	.21 (.018)	.28 (.022)	.41 (.015)	.41 (.017)	.41 (.016)	.41 (.020)	.11 (.022)	.07 (.013)	.40 (.020)	.40 (.021)	.39 (.022)
$\rho_{\Delta c,y}(1)$	.27 (.012)	.24 (.021)	.22 (.017)	.27 (.012)	.28 (.013)	.28 (.012)	.28 (.013)	.16 (.017)	.07 (.014)	.27 (.012)	.27 (.013)	.27 (.014)
$\rho_{\Delta c,y}(2)$	.23 (.011)	.20 (.018)	.18 (.015)	.23 (.011)	.23 (.012)	.23 (.011)	.23 (.012)	.13 (.015)	.05 (.013)	.23 (.012)	.22 (.012)	.22 (.013)
$\rho_{y,r}(1)$	.48 (.028)	.46 (.040)	.42 (.026)	.49 (.028)	.49 (.028)	.49 (.028)	.47 (.034)	.35 (.025)	.12 (.053)	.49 (.033)	.49 (.033)	.49 (.034)
$\rho_{y,r}(2)$	.47 (.027)	.45 (.039)	.41 (.026)	.48 (.027)	.48 (.027)	.48 (.027)	.46 (.033)	.34 (.024)	.12 (.052)	.47 (.032)	.47 (.031)	.47 (.033)
$\rho_{y,r}(3)$	.45 (.027)	.43 (.039)	.40 (.026)	.46 (.027)	.46 (.027)	.46 (.027)	.44 (.032)	.33 (.023)	.12 (.051)	.46 (.030)	.46 (.030)	.46 (.032)

*See notes to Tables 4 and 5.

Table 7

Stochastic Properties: Three State, $\sigma_\varepsilon = .01$, $\tau = 3.0$ Economy*

Statistic	DP	Log	Lin	Statistics	DP	Log	Lin	Statistic	DP	Log	Lin
C	3.94	3.94	3.94	σ_c/σ_y	.76	.76	.76	$\rho_{\Delta c,r}(1)$	.107	.090	.090
(3.94)	(.020)	(.020)	(.020)		(.026)	(.026)	(.026)		(.012)	(.011)	(.012)
Y	3.94	3.94	3.94	σ_{dk}/σ_y	.49	.49	.48	$\rho_{\Delta c,r}(2)$	.053	.087	.087
(3.94)	(.020)	(.020)	(.020)		(.022)	(.022)	(.022)		(.011)	(.011)	(.012)
K	63.97	63.72	63.61	σ_r/σ_y	.006	.005	.005	$\rho_{\Delta c,\Delta c}(1)$	.026	.028	.028
(63.69)	(.59)	(.59)	(.59)		(.0001)	(.0001)	(.0001)		(.011)	(.011)	(.011)
K/Y	16.22	16.17	16.16	σ_y	.170	.169	.169	$\rho_{\Delta c,\Delta c}(2)$	.024	.023	.024
(16.17)	(.059)	(.069)	(.069)		(.010)	(.010)	(.010)		(.013)	(.014)	(.013)
MP_k	.0204	.0204	.0205	$\rho_{r,c}(0)$	-.40	-.47	-.48	$\rho_{\Delta c,y}(1)$	-.005	-.005	-.005
(.0204)	(.86E-4)	(.87E-4)	(.88E-4)		(.049)	(.057)	(.062)		(.008)	(.009)	(.009)
R	.0204	.0204	.0204	$\rho_{dk,c}(0)$	.26	.26	.26	$\rho_{\Delta c,y}(2)$	-.009	-.009	-.009
(.0204)	(.88E-4)	(.87E-4)	(.89E-4)		(.022)	(.022)	(.022)		(.008)	(.008)	(.009)
				$\rho_{r,\Delta c}(0)$	.25	.30	.30	$\rho_{y,r}(1)$	-.031	-.035	-.041
					(.012)	(.015)	(.017)		(.069)	(.082)	(.089)
								$\rho_{y,r}(2)$	-.036	-.042	-.047
									(.068)	(.081)	(.088)

*The model economy underlying the simulations in this table is the same as the "three-state low variance" economy in Tables 4-6, with the exception that here, $\tau = 3.0$. See notes to Tables 4 and 5 for further information on notation and details of the simulation experiments.

Table 8: Cost, in Percent of K_{t-1}, of Using LQ Decision Rules[1]
(How Much an LQ Planner Would Pay to Learn the DP Rule)

Panel A: $\sigma_\epsilon = .01$, 2–state Markov chain, $\tau = .5$

	x_t=−.032		x_t=.032	
K_{t-1}	Log	Lin	Log	Lin
60.32	0.0	0.0	0.0	0.0
62.00	0.0	0.0	0.0	0.0
63.69	0.0	0.0	0.0	0.0
65.46	0.0	0.0	0.0	0.0
67.23	0.0	0.0	0.0	0.0

Panel B: $\sigma_\epsilon = .10$, 2–state Markov chain, $\tau = .5$

	x_t=−.32		x_t=0.32	
K_{t-1}	Log	Lin	Log	Lin
36.78	0.3	0.3	0.6	0.3
50.24	0.2	0.3	0.2	0.2
63.69	0.1	0.2	0.1	0.2
86.19	0.2	0.2	0.1	0.2
108.69	0.3	0.1	0.2	0.2

Panel C: $\sigma_\epsilon = .01$, 3–state Markov chain, $\tau = .5$

	x_t=−.06		x_t=0.00		x_t=0.06	
K_{t-1}	Log	Lin	Log	Lin	Log	Lin
57.96	0.0	0.0	0.0	0.0	0.0	0.0
60.82	0.0	0.0	0.0	0.0	0.0	0.0
63.69	0.0	0.0	0.0	0.0	0.0	0.0
66.82	0.0	0.0	0.0	0.0	0.0	0.0
69.96	0.0	0.0	0.0	0.0	0.0	0.0

Panel D: $\sigma_\epsilon = .10$, 3–state Markov chain, $\tau = .5$

	x_t=−.55		x_t=0.00		x_t=0.55	
K_{t-1}	Log	Lin	Log	Lin	Log	Lin
24.60	1.0	2.6	1.1	1.2	5.1	1.5
44.14	0.3	2.0	0.3	0.8	1.1	1.2
63.69	0.3	1.2	0.2	0.6	0.4	1.1
112.26	0.7	0.5	0.2	0.4	0.2	0.8
160.84	1.4	0.3	0.4	0.3	0.3	0.6

[1]Let $v^* = v_{LQ}(K_{t-1},x_t)$, for given K_{t-1}, x_t, where LQ refers either to log–linear LQ or linear LQ, as indicated in the column headings. Let K^* be such that $v_{DP}(K^*,x_t) = v^*$. The entries in the table are $[(K_{t-1}-K^*)/K_{t-1}]\times 100$.

Appendix 1: Value Function Iteration.

This section describes in detail the value iteration method used to solve the dynamic programming problem, (3.1). Although the example in the text is a simple one, the discussion that follows is quite general. In particular, the discussion generalizes trivially to the case where K and x are vectors. In addition, the absence of choice variables (like hours worked) which affect u but not v directly does not limit the generality of what follows, since one is free to think of u as the indirect utility function obtained after maximizing out such control variables. Finally, Christiano and Fitzgerald (1983) show how a problem with multiple decisions and exotic information constraints also fits into the following framework. The first part of this appendix discusses several value iteration methods used to solve the discretized model. The second part of this section reports the computing times used by each.

A.1 Three Value Iteration Methods

To apply value iteration methods, the optimization problem first has to be expressed as a dynamic programming problem. The notation is simplified if (3.1) is expressed without time subscripts:

$$(A.1) \qquad v(K,x) = \max_{K' \in \Lambda(K,x)} \{u(K,K',x) + \beta E_{x'} [v(K',x') | x]\}.$$

The relationship between the variables in (A.1) and their time subscripted counterparts is as follows. The variables K and x correspond to K_{t-1} and x_t, respectively. The variables K′ and x′ correspond to K_t and x_{t+1}, respectively. The expectation in (A.1) is over values of x′ conditional on x and the Markov chain model in (2.3). Also, Λ denotes the feasible set of possible choices of K′, given the

state, K, x. It is formally defined in (3.2), which is reproduced here for convenience:

$$\text{(A.2)} \qquad \Lambda(K,x) \equiv \{K' \in \mathcal{K}: \exp(x)K^{\alpha} + (1-\delta)K - K' \geq 0\}.$$

The set $\mathcal{K}$ is the capital grid discussed in section 3.

Because $\mathcal{K} \times \mathcal{X}$ is a discrete set of $m = n_x \times n_k$ points, it follows that v can be represented as a point in $\mathbb{R}^m$. Here, n_k and n_x denote the number of elements in $\mathcal{K}$ and $\mathcal{X}$, respectively. In particular, let

$$\text{(A.3)} \qquad \begin{aligned} s_{[(j-1)n_x+i]} &= (\mathcal{K}_j, \mathcal{X}_i)', \; i = 1,\ldots,n_x, \; j = 1,\ldots,n_k, \\ S &= (s_1', \ldots, s_m')'. \end{aligned}$$

Then S is a $2m \times 1$ vector enumerating all the possible states. The $m \times 1$ vector v is:

$$\text{(A.4)} \qquad v = (v(s_1), \ldots, v(s_m))'.$$

The expression on the right hand side of (A.1) defines a function, T, from $\mathbb{R}^m$ into itself. From this perspective, the optimization problem is solved by first finding a $v \in \mathbb{R}^m$ such that

$$\text{(A.5)} \qquad v = T(v).$$

The existence and uniqueness of a $v \in \mathbb{R}^m$ satisfying (A.5) is an implication of the contraction mapping theorem (Sargent [1987,p.343]) and the fact that T is a contraction mapping. The latter in turn is readily established by confirming that T satisfies Blackwell's sufficient conditions to be a contraction (Sargent [1987,p.344–45]). After finding v, the decision rule sought is given by:

$$(A.6) \qquad f_{DP}(K,x) = \underset{K' \in A(K,x)}{\operatorname{argmax}} \{u(K,K',x) + \beta E_{x'}[v(K',x')|x]\}.$$

Value function iteration methods for finding a point v satisfying (A.5) share the characteristic that starting with an initial $v_0 \in \mathbb{R}^m$, they compute a sequence $v_j \in \mathbb{R}^m$ and set $v = \lim_{j\to\infty} v_j$. Following are three such methods.

Standard Value Iterations

The simplest value iteration method is *standard value iteration.* It generates the sequence $\{v_j\}$ by iterating on T: $v_j = T(v_{j-1})$, $j = 1,2,3,\ldots$, usually with $v_0 = 0$. Convergence of this sequence to v in the sup norm sense is guaranteed by the contraction mapping theorem. The method is also called the method of *successive approximation* (Bertsekas [1976, p.237]) or *contraction iterations* (Rust [1987,1988a,1988b].) It is useful to express the $j+1^{st}$ iterate explicitly as a function of v_j. First, let the m×m matrix G_j denote the state transition probability matrix implied by π and the decision rule, $f^j_{DP}(K,x)$ computed by replacing v by v_j in (A.6). Specifically, the ℓ,ν^{th} element of G_j is the probability that $(K',x') = s_\nu$ given $(K,x) = s_\ell$. Also, let u_j be the m×1 vector $u_j = [u(s_1,f^j_{DP}(s_1)),u(s_2,f^j_{DP}(s_2)),\ldots,u(s_m,f^j_{DP}(s_m))]'$. In this notation, it is easy to confirm:

$$(A.7) \qquad T(v_j) = u_j + \beta G_j v_j.$$

The standard value iteration method algorithm can be slow to converge, and this has stimulated interest in computationally more efficient algorithms.

One alternative to iterating on T by standard value function iterations uses Newton's method. Given v_j, this method equates v_{j+1} with the fixed point of the linear Taylor series expansion of T about $v = v_j$. Specifically, the linearized T function is $T(v_j) + T'(v_j)(x - v_j)$, where $T'(v_j)$ is the derivative of T(x) with respect to x, evaluated at $x = v_j$. Since small perturbations in v_j have no effect on f^j_{DP}, and therefore u_j and G_j, it follows from (A.7) that,

$$(A.8) \qquad T'(v_j) = \beta G_j.$$

Then, v_{j+1} is by assumption the fixed point of this function, i.e., $v_{j+1} = v_j + [I - T'(v_j)]^{-1}[T(v_j) - v_j]$. For reasons made clear below, I use T_∞ to signify this operator:

$$(A.9) \qquad \begin{aligned} T_\infty(v_j) &\equiv v_j + [I - T'(v_j)]^{-1}[T(v_j) - v_j] \\ &= [I - \beta G_j]^{-1} u_j. \end{aligned}$$

The expression after the second equality was obtained by substituting $T'(v_j)$ and $T(v_j)$ out of the expression after the first equality using (A.8) and (A.7), respectively.[9] The existence and uniqueness of $T_\infty(v_j)$ in (A.9) follows from the facts that $0 < \beta < 1$ and that the modulus of the largest eigenvalue of G_j is unity. The latter is an implication of the fact that G_j is a transition probability matrix (see Nobel [1969,p.458].) I elaborate on this in the proof to Proposition 3 below.

Because in a discrete state space there exists only a finite set of feasible

[9]Rust (1987,1988a,1988b) iterates on value functions using the operator to the right of the identity sign in (A.9) and calls the method *Newton–Kantorovich iteration*, whereas one of Bertsekas' suggested value iteration methods is defined by the operator to the right of the equality sign in (A.9). Bertsekas (1976,p.246) calls this value function iteration method the *policy iteration algorithm*. The policy iteration algorithm is also known as the *Howard policy improvement algorithm* (see Sargent [1987,p.47].) The identity in equation (A.9) establishes the mathematical equivalence of these algorithms.

decision rules, it follows that the set of objects G_j and u_j consistent with feasibility is also finite. As a result, the sequence of Newton iterates, $\{v_j, j = 1,\dots\}$, is contained in a finite set of points in $\mathbb{R}^m$. Moreover, a simple argument (see Bertsekas [1976,p.246]) establishes that $v_{j+1} \geq v_j$ for $j \geq 1$. Consequently, Newton iterations converge in a finite number of steps to a v such that $v = T_\infty(v)$. This in turn implies that $v = T(v)$, and is therefore the object sought (to see this, just set $T_\infty(v_j) = v_j$ in [3.9].)

There are two difficulties with the Newton value iteration method: one is only apparent and the other real. Although the notation used in (A.9) suggests that the method requires a huge amount of computer storage, in fact $[I - T'(v_j)]$ is composed mostly of zeroes, and these do not need to be stored. A more substantive problem with the method is the requirement of solving the m equations $[I - \beta G_j]T_\infty(v_j) = u_j$ in the m unknown elements of $T_\infty(v_j)$. Unless the structure of the optimization problem is such that some recursive algorithm for doing this rapidly is available, then direct application of the procedure is computationally prohibitive. We therefore seek computationally efficient ways to approximate the solution to these equations. The following value function iteration method is one way to do this.

Hybrid Value Function Iterations

Denote the value of using the feasible policy f for one period given that next period's state variables are valued according to the value function x by $T_f(x)$. By construction, $T_{f^j_{DP}}(v_j) \equiv T(v_j)$. Let $T^2_{f^j_{DP}}(v_j) \equiv T_{f^j_{DP}}[T_{f^j_{DP}}(v_j)]$ be the value of following policy f^j_{DP} for two periods given that the subsequent period's state is valued according to v_j, and define $T^p_{f^j_{DP}}(v_j)$ similarly, for $p = 2,3,\dots$. The third value function iteration method computes v_{j+1} from v_j as follows:

(A.10) $$v_{j+1} = T_p(v_j) \equiv T^p_{f_{DP}}(v_j).$$

I call the value iteration method which uses T_p *hybrid value function iteration.* When $p = 1$, then this method reduces to the standard value function iteration method, i.e., $T_1 \equiv T$. In addition, for sufficiently large p, hybrid value iterations approximate Newton value iterations arbitrarily well. That is

Proposition 3: for any $v_j \in \mathbb{R}^m$

(A.11) $$T_\infty(v_j) = \lim_{p\to\infty} T_p(v_j),$$

where T_∞ and T_p are defined in (A.9) and (A.10), respectively.

Proof:

It is easy to confirm, using the fact that the modulus of the maximal root of βG_j is β, that $T_{f_{DP}}$ satisfies Blackwell's sufficient conditions to be a contraction mapping. Consequently, by the contraction mapping theorem, there exists a unique $x \in \mathbb{R}^m$ with the properties:

(A.12) $$x = T_{f_{DP}}(x)$$

(A.13) $$x = \lim_{\ell\to\infty} T^\ell_{f_{DP}}(x_0) \text{ for any } x_0 \in \mathbb{R}^m.$$

Since (A.13) is valid for $x_0 = v_j$, it can be rewritten:

(A.14) $$x = \lim_{p\to\infty} T^p(v_j).$$

Writing (A.12) out explicitly,

(A.15) $$x = u_j + \beta G_j x.$$

Since the $x \in \mathbb{R}^m$ that solves (A.15) is unique and (A.9) defines $T_\infty(v_j)$ as a solution to (A.15) it follows that

(A.16) $$x = T_\infty(v_j).$$

Equations (A.14) and (A.16) imply (A.11). Q.E.D.

What hybrid policy iteration does, essentially, is to approximates the matrix $[I - T'(v_j)]^{-1}$ by its p^{th} order series expansion: $I + T'(v_j) + T'(v_j)^2 + ... + T'(v_j)^p$.

A.2 Some Computational Results.

Table A1 reports the results of some experiments comparing the efficiency of the standard, Newton and hybrid value function iteration methods. The calculations were done for the three state, low variance, $\tau = .5$ version of the model in the text. Since $n_k = 20{,}000$, $n_x = 3$, the state space is composed of $m = 60{,}000$ points. The convergence criterion used in the calculations in Table A1 is that the maximum percent difference between v_j and v_{j-1} be less than $.1 \times 10^{-7}$ and there be no difference between f^j_{DP} and f^{j-1}_{DP}. This criterion is tighter than the one underlying the calculations underlying Table 1, which is reported in section 4.b.

The first row in Table A1 reports results for solving the model by standard value iterations, i.e., with $p = 1$ throughout the calculations. Convergence occured at $j = 947$ steps in 414.11 central processor unit (CPU) minutes. Upon inspecting the output I found that the decision rule had actually stopped changing after the $j = 300$ step, so that after this, the algorithm was essentially iterating on $T_{f^{300}_{DP}}$. This suggests that setting $p = \infty$ later in the calculations would improve things. I modified the routine so that it triggered into Newton iterations (i.e., $p = \infty$) when the decision rule failed to change at 10% of the points in the state space. The results of this appear in row 2 of Table A1. That row shows that triggering into Newton iterations at the end reduced the number of steps required for convergence by a factor of about 9. However, the amount of CPU time used only fell by about 25 percent, reflecting the greater cost per step of doing Newton rather than standard iterations. Row 3 reports the consequences of setting $p = 10$ at the start of the iterations, and then triggering into Newton iterations when the

algorithm is close to convergence. This resulted in a further slight reduction in CPU time of roughly 7%. The bottom row in Table A1 shows the effect of setting p = 10 thoughout the calculations, and not triggering into Newton iterations at the end. That results in about the same computer time when the trigger is applied (row 3).

I conclude that hybrid policy iterations with p = 10 are an improvement over standard value iterations, and that triggering into Newton iterations at the end makes little difference. Although the 25% improvement in CPU time relative to standard value iterations is useful, the results using hybrid value iteration obtained here are not as dramatic as those reported in Christiano and Fitzgerald (1988). In that paper, hybrid value iterations led to more than a 10 fold reduction in CPU time and triggering into Newton iterations at the end reduced CPU time by another factor of 3. The difference in the results probably reflects that the maximization in (A.1) is very inexpensive in the example of this paper, whereas in Christiano and Fitzgerald (1988) it is more costly because four control variables rather than one are involved.

Table A1

Computational Costs of Alternative Versions of Hybrid Policy Iterations[1]

p[2]	Trigger[3]	Steps[4]	CPU[5] Time
1	no	947	414.11
1	6000	165	356.57
10	6000	118	332.77
10	no	101	324.78

[1]Calculations based solving the three state, low variance, $\tau = .5$ model using a tighter convergence criterion than the calculations underlying those in Table 1. The convergence criterion underlying the calculations here was that the maximum percent difference between v_j and v_{j-1} be less than $.1 \times 10^{-7}$ and that there be no change in the decision rule.

[2]Value of p in hybrid value function iteration for j = 1, 2, 3, ..., until the decision rule changes at less than *trigger* points in the state space, which is itself composed of 60,000 points. After this, $p = \infty$. When p = 1 the method corresponds to standard value function iterations. When $p = \infty$ the method represents Newton value function iterations.

[3]A no in this column means no trigger was used.

[4]Number of value function iteration steps.

[5]Central Processing Unit minutes on the Federal Reserve Bank of Minneapolis' Amdahl dual 580 mainframe computer.

References

Altug, Sumru, 1986, "Time to Build and Aggregate Fluctuations: Some New Evidence," manuscript, Federal Reserve Bank of Minneapolis.

Baxter, Marianne, 1988, "Approximating Suboptimal Dynamic Equilibria: An Euler Equation Approach," Rochester Center for Economic Research Working Paper no. 139, June.

Bertsekas, Dimitri, 1976, Dynamic Programming and Stochastic Control, Academic Press.

Brock, William A., and Leonard J. Mirman, 1972, "Optimal Economic Growth and Uncertainty: The Discounted Case," Journal of Economic Theory 4,479–513.

Christiano, Lawrence J., 1986, "On the Accuracy of Linear Quadratic Approximations: An Example," Federal Reserve Bank of Minneapolis.

Christiano, Lawrence J., 1987a, "Intertemporal Substitution and the Smoothness of Consumption," Manuscript, Federal Reserve Bank of Minneapolis.

Christiano, Lawrence J., 1987b, "Dynamic Properties of Two Approximate Solutions to a Particular Growth Model," Research Department Working Paper no. 338, Federal Reserve Bank of Minneapolis.

Christiano, Lawrence J., 1987c, "Is Consumption Insufficiently Sensitive to Innovations in Income?", American Economic Review Papers and Proceedings, vol. 77, no. 2, May.

Christiano, Lawrence J., 1988, "Why Does Inventory Investment Fluctuate So Much?," Journal of Monetary Economics, 21, 247–280.

Christiano, Lawrence J. and Martin Eichenbaum, 1988, "Is Theory Really Ahead of Measurement? Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations", Working Paper 412, Federal Reserve Bank of Minneapolis, August.

Christiano, Lawrence and Terry Fitzgerald, 1988, "The Magnitude of the Speculative Motive for Holding Inventories in a Real Business Cycle Model," manuscript, Federal Reserve Bank of Minneapolis, October.

Cooley, Thomas F., and Gary D. Hansen, 1988, "The Inflation Tax in a Real Business Cycle Model," Department of Economics, UCLA Working paper no. 496.

Coleman, T., 1984, "Essays on Aggregate Labor Market Business Cycle Fluctuations," Ph.D. thesis, University of Chicago, 1984.

Danthine, Jean–Pierre, John B. Donaldson and Rajnish Mehra, 1988, "On Some Computational Aspects of Real Business Cycle Theory," Sloan School of Management working paper number 1984–88, Massachusetts Institute of Technology, June.

Danthine, Jean–Pierre, John B. Donaldson and Lance Smith, 1987, "On the Superneutrality of Money in a Stochastic Dynamic Macroeconomic Model," Journal of Monetary Economics 20, no.3, December.

Hansen, Gary D., 1985, "Indivisible Labor and the Business Cycle," Journal of Monetary Economics 16, 309–327

Hansen, Gary D., and Thomas Sargent, 1988, Journal of Monetary Economics 21

King, Robert G., Charles Plosser and Sergio Rebelo, 1988, Journal of Monetary Economics 21.

Kydland, Finn, and Edward Prescott, 1980, "A Competitive Theory of Fluctuations and the Feasibility and Desirability of Stabilization Policy," in Stanley Fischer, ed., Rational Expectations and Economic Policy (University of Chicago Press, Chicago, Illinois).

Kydland, Finn, and Edward Prescott, 1982, "Time to Build and Aggregate Fluctuations," Econometrica 50, 1345–1370.

Long, John B., and Charles I. Plosser, 1982, "Real Business Cycles," Journal of Political Economy.

Mehra, Raj and Edward C. Prescott, 1985,"The Equity Premium: A Puzzle," Journal of Monetary Economics 15, pp. 145–161.

Noble, Ben, 1976, Applied Linear Algebra, Prentice–Hall, Englewood Cliffs, New Jersey.

Prescott, Edward C., 1986, "Theory Ahead of Business Cycle Measurement," Federal Reserve Bank of Minneapolis Quarterly Review, fall.

Rust, John, 1987, "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," Econometrica.

Rust, John, 1988a,"Maximum Likelihood Estimation of Discrete Control Processes," SIAM Journal of Control and Optimization, vol.26,no.5, November.

Rust, John, 1988b, "A Dynamic Programming Model of Retirement Behavior," in David Wise, ed., The Economics of Aging, National Bureau of Economic Research, University of Chicago Press.

Sargent, Thomas J, 1979, "Tobin's q and the Rate of Investment in Equilibrium," Federal Reserve Bank of Minneapolis Staff Report 40.

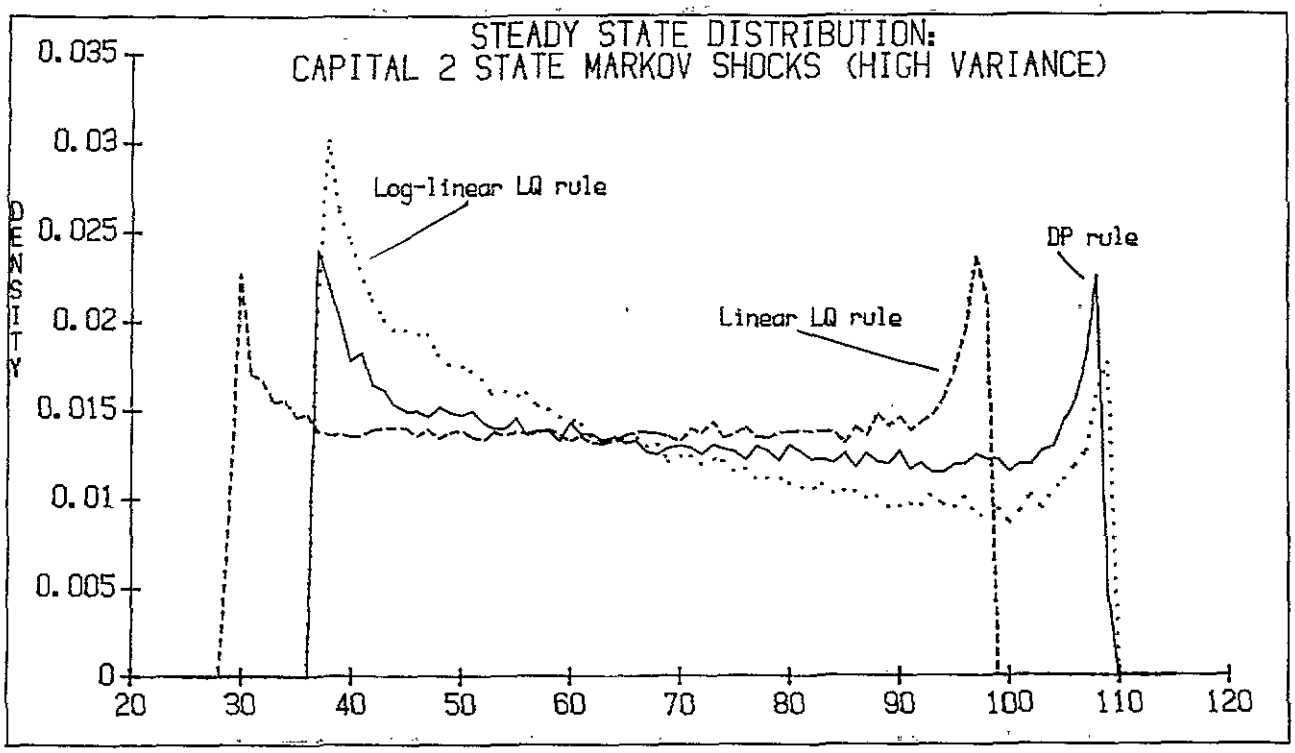

FIGURE 1a

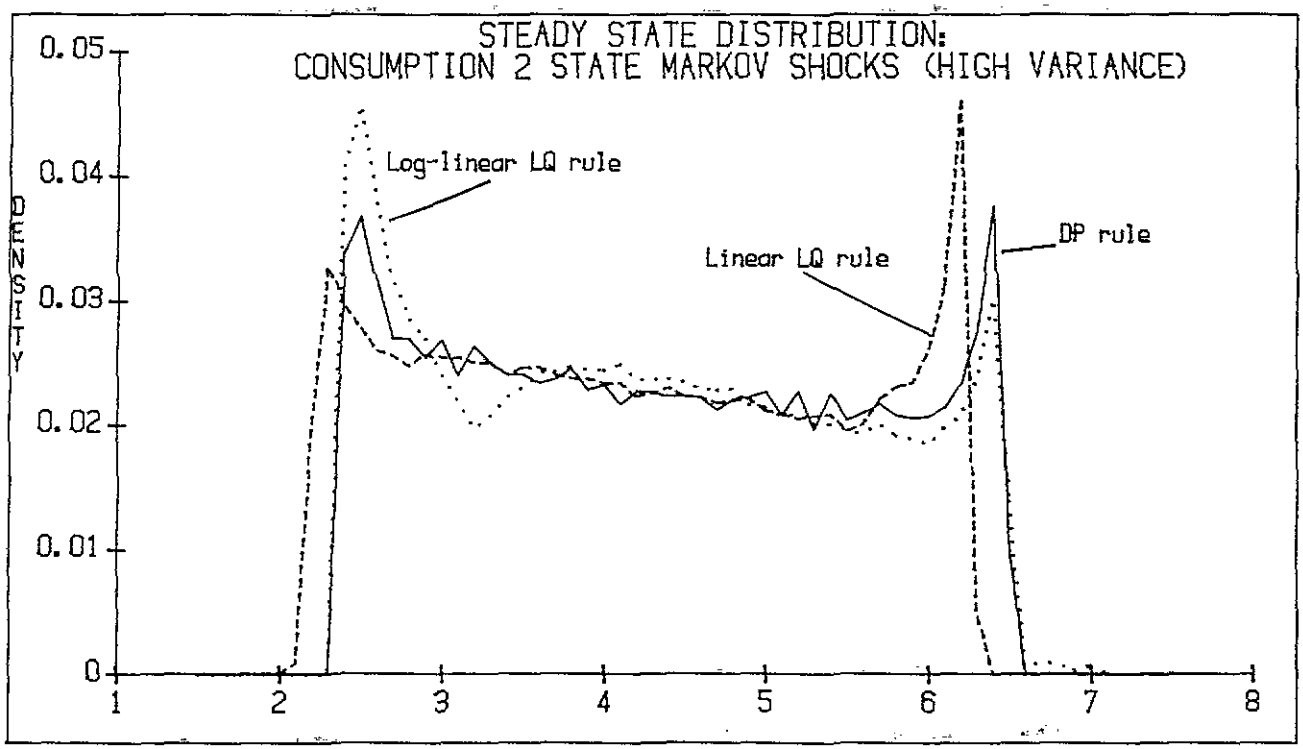

FIGURE 1b

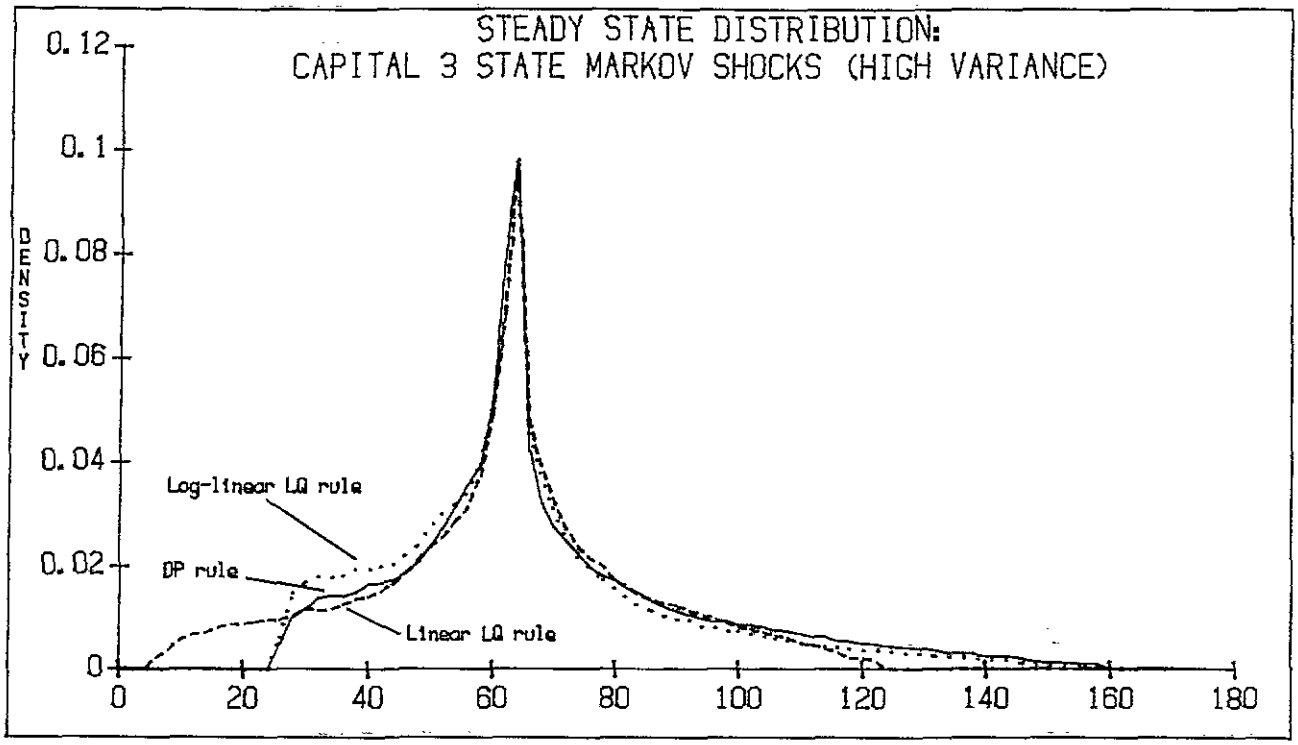

FIGURE 2a

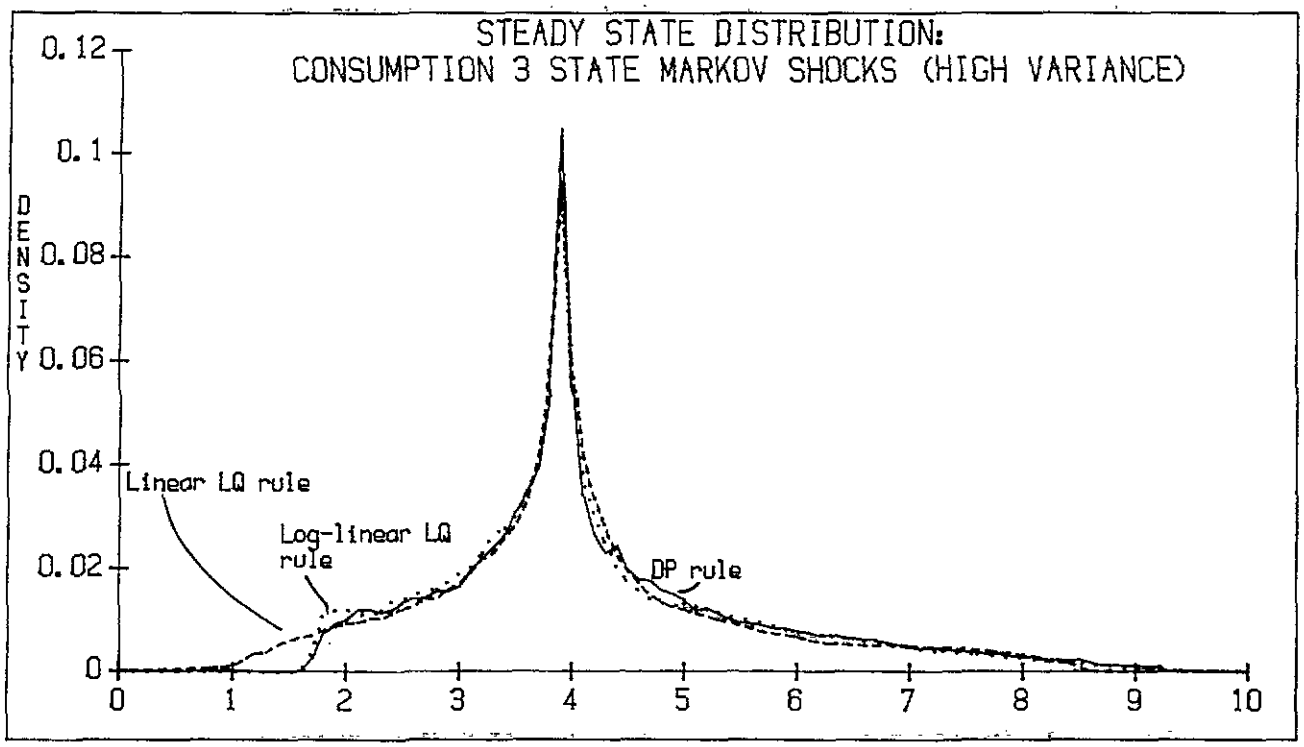

FIGURE 2b