UNIVERSITAT POMPEU FABRA

Solving Non-Linear Stochastic Models by Parameterizing Expectations: An Application to Asset Pricing with Production

Albert Marcet

Economics working paper 5.July 1991

Solving Non-Linear Stochastic Models by Parameterizing Expectations: An Application to Asset Pricing with Production

by

Albert Marcet[1]

Universitat Pompeu Fabra
GSIA, Carnegie Mellon University

July 1991

---

[1]Sections 1 and 2 are a rewrite of my earlier paper "Solving Non-Linear Dynamic Models by Parameterizing Expectations" [1988]. I am indebted to Wouter den Haan, Joan Ketterer, David Marshall and Guido Tabellini. I also thank the participants of a seminar I gave at Stanford University in January 1989: Ken Judd, Rody Manuelli, Ramon Marimon, John O'Brien, Richard Rogerson, Tom Sargent and Ken Singleton. The meetings of the NBER Non-Linear Rational Expectations Modelling Group were extremely helpful; the discussions in that group had a large influence in this paper. All errors are my own. This research was supported by a grant from the National Science Foundation.

UNIVERSITAT POMPEU FABRA
BIBLIOTECA
1001971371

Summary

A new algorithm called the parameterized expectations approach (PEA) for solving dynamic stochastic models under rational expectations is developped and its advantages and disadvantages are discussed. This algorithm can, in principle, approximate the true equilibrium arbitrarily well. Also, this algorithm works from the Euler equations, so that the equilibrium does not have to be cast in the form of a planner's problem. Monte-Carlo integration and the absence of grids on the state variables, cause the computation costs not to go up exponentially when the number of state variables or the exogenous shocks in the economy increase.

As an application we analyze an asset pricing model with endogenous production. We analyze its implications for time dependence of volatility of stock returns and the term structure of interest rates. We argue that this model can generate hump-shaped term structures.

keywords: numerical algorithm, Monte-Carlo integration, rational expectations, bond prices.

Address: Albert Marcet
Facultat d'Economiques
Universitat Pompeu Fabra
Balmes, 132
08008, Barcelona
Spain

Introduction

During the last decade, the use of dynamic, stochastic models has extended to all fields of economics. Perhaps the main limitation of this class of models is that they are very difficult to solve; few of them can be solved analytically[2] and only after making very strong assumptions.

Given these difficulties, and given the progress that computer hardware has experienced in the last few years, it is worthwhile to put some effort on methods for solving these models numerically. With numerical solutions it is possible to simulate the models and perform exercises of theoretical interest (how does the model respond to a change in the environment?, to a change in policy?, does the model reproduce some stylized facts?), as well as empirical exercises (estimation by simulation or calibration).

The main object of this paper is to introduce the parameterized expectations approach (PEA) for calculating numerical solutions to stochastic non-linear models with rational expectations; we discuss its flexibility, its advantages and disadvantages. Despite its short life, this procedure has been applied successfully to a number of models: Marshall [1988] uses it to solve a model of money with transactions costs, Den Haan [1988] solves a model of money with a shopping-time technology and compares the equilibrium under the optimal monetary policy with some suboptimal policy rules, Marcet and Singleton [1989] show how to solve asset pricing models with heterogeneous agents and credit constraints, Ketterer and Marcet [1989] apply it to study the effect of introducing options in financial markets with incomplete markets; Marcet and Marimon [1989] solve a stochastic growth model with incentive compatibility and participation constraints, and den Haan [1991] applies it to study the term structure of interest rates in a model with money and endogenous production and Rojas [1991] solves for the optimal taxation problem in a model with private and public capital.

The basic idea in PEA is to approximate the conditional expectations that appear in Euler equations by parameterizing them with flexible functional forms; we then iterate on the parameters of these functions until the expectations are good predictors by looking at simulations done with several parameter values. This will usually

---

[2]Some examples are the linear-quadratic case, some models with exponential utility, the Brock & Mirman version of the neoclassical growth model, or the asset pricing model with discrete dividends (Mehra and Prescott [1985]).

involve numerically calculating the fixed point to a certain mapping in order to guarantee that, given that agents use these expectations, these are in fact the best predictors of the future. We will discuss techniques based on homotopy that can be used in calculating the fixed point.

Because in PEA we check if the expectations are good predictors along simulations, we are ensured that the approximation is good in those intervals of the state variables that happen more often. When an algorithm gives more importance to those values of the variables that happen more often it said that this algorithm performs "endogenous oversampling"; by its nature, PEA does endogenous oversampling and Monte-Carlo integration at the same time.

The parameterized expectations approach can, in principle, approximate the equilibrium with arbitrary accuracy by refining the degree of approximation; for example, if the conditional expectation is parameterized as a polynomial, one can use a polynomial degree high enough to obtain arbitrary accuracy[3]. Perhaps its main advantage, however, resides in its great flexibiliy that makes it applicable to a large number of models with moderate computational costs.

As an application with independent interest, we study an asset pricing model with endogenous production and capital accumulation as in Brock [1982]. This is an example of a relatively simple model that is very difficult to characterize analytically but that can be easily analyzed by simulation[4]. We show that a very simple model with endogenous production generates a very rich pattern of covariances between asset prices and interest rates that would be much harder to capture in a model with exogenous production and it improves the performance of the model in some aspects. In particular it generates long-term dependence between asset prices and the real part of the economy, and it creates a humped-shaped term structure of interest rates; we also argue that, even though it displays higher variance of stock returns in recessions as discussed by Rowenhorst [1991], this result would happen in most asset pricing models. Finally, it has been argued that one of the reasons for the equity premium puzzle in the paper by Mehra and Prescott [1985] is that they force the volatility of stock returns to be very small compared with the true data, due to the fact that dividends are perfectly correlated with total production; the model with endogenous production provides a natural

---

[3] Den Haan and Marcet [1989] discuss a method for evaluating the accuracy of numerical solutions to dynamic models. They also evaluate the accuracy of PEA in several models.

[4] Balduzzi [1991] studies a version of this model with an analytic solution, but he has to assume no depretiation of capital. Rowenhorst [1991] and den Haan [1991] study a similar model by simulation.

way of making stock returns highly volatile without introducing many parameters in the model, since payments of capital are decided endogenously; we show that, even when the stock returns are made as volatile as in the US data in this case, this model still does not solve the equity premium puzzle.

Much progress has been made recently on the subject of numerical solution procedures. Some algorithms are based on discrete state space techniques used by Miller [1984], Pakes [1986], Rust [1987], Tauchen [1986] and Wolpin [1984] and Baxter [1988]. Also, the extended path solution procedure of Fair and Taylor [1983], the Euler equation approach of Labadie [1986], the linear-quadratic approximation as in Kydland and Prescott [1982], the backwards solution technique of Sims [1986], Novales [89] and Ingram [1986] and the Euler equation approaches of Coleman [1988], Judd [1989] and Baxter [1988] are available to solve dynamic stochastic models[5]. The ideas in this paper are very related to other solution procedures that existed before the earlier version; in particular, the relationship to the backwards solution method of Sims and Novales, to the iterations performed in Townsend [1983] and to Marcet and Sargent [1989b] will be discussed.

The traditional techniques based on dynamic programming, namely, the linear quadratic approximation and the value function iterations, have provided many useful applications in economics, but there are many models of interest that can not be addressed with these numerical techniques. The value function iterations need a large grid on the state space to be imposed , so that the computational constraints of existing hardware are binding, and the linear quadratic approximation has the problem of not being applicable in models with inequality constraints, and it does not provide an arbitrarily good approximation. Also, these techniques are designed to solve a planner's problem, and sometimes it is difficult to cast equilibrium models in planner's problems.

The parameterized expectations approach can, in principle, give an arbitrarily good approximation to the rational expectations equilibrium, it is flexible enough so that it can be applied to a large class of models, and it can solve suboptimal equilibria "directly", without having to characterize the equilibrium as the solution to a planner's problem. Even though the techniques described in the present paper seem quite useful it is certainly not the case that they uniformly dominate the dynamic programming methods. For example, there are models where the linear quadratic approximation can be applied and where it is a fairly good approximation; since PEA is more computationally intensive, in these cases one would want to use the LQ approximation. In models where the decision variables are

---

[5] See the Taylor and Uhlig [1990] paper for a comparison of some of these techniques in solving a simple growth model.

discrete and they can take on very few values, iterations on the value function are likely to be a a better approach. The recent techniques of Coleman and Judd can be applied in similar circumstances as PEA. We will discuss the advantages and disadvantages of PEA in front of the alternatives in the text.[6]

We propose several numerical algorithms to calculate a certain fixed point involved in the solution. The algorithms we suggest in this paper converge to the fixed point if the model at hand is stable under a particular kind of non-linear least squares learning mechanism. This provides some peace of mind in that those models that are stable under learning can be calculated with these algorithms. Also, for models with multiple equilibria it is often the case that only one of them is stable under learning; for these models, the algorithms proposed here would find the stable equilibrium, and the researcher does not need to worry about the other equilibria. Finally, one should keep in mind that stability of least squares learning algorithms is a fairly mild requirement, as most dynamic models that have been analyzed in the learning literature turn out to be locally stable.[7]

We start by discussing the general idea of the algorithm, and move gradually to more systematic applications. In section 1, we describe the parameterized expectations approach, its motivation, we show how it can be applied to several models. Section 2 discusses numerical algorithms for solving the fixed point problem involved in the parameterized expectations approach, and how to use the homotopy approach for this purpose. In section 3 we solve an asset pricing model with endogenous production by PEA; we discuss some practical considerations that arise in solving a simple growth model, and we discuss the speed of the algorithm; we study the implications of the model for the term structure of interest rates, and the volatility of stock returns. Section 4 contains some conclusions and suggestions for extensions.

---

[6]For a review of simulation algorithms of dynamic stochastic models and their applications see Marcet [1991].

[7] See Marcet and Sargent [89a], [89b] for some of these results and for other references

1. The Parameterized Expectations Approach

The problem at hand is to find numerically the solution to the endogenous variables of a stochastic, dynamic model as functions of the exogenous variables and for given parameter values of the parameters of the model.

More specifically, consider a dynamic model with rational expectations, where the economy is described by a vector $z_t \in R^n$. Each period there are s exogenously given i.i.d. shocks $u_t \in R^s$, and agents have an information set Ω_t. There is a large class of models where the equilibrium process for $\{ z_t \}$ satisfies

$$(1.1) \qquad g(E [\phi(z_{t+1}) \mid \Omega_t], z_t, z_{t-1}, u_t) = 0 \qquad \text{for all } t,$$

for known functions $\phi:R^n \to R^m$ and $g:R^{m+2n+s} \to R^n$; here n is the number of variables in the vector z_t, m is the number of conditional expectations involved in the system (1.1), and s is the number of i.i.d. shocks in the system. Here, the functions g and ϕ will depend on the parameters of the model; This dependence is left implicit for most of the paper. Equations that will typically be part of the system (1.1) will be Euler equations, resource contraints, laws of motion of exogenous processes, market clearing conditions, implementability constraints, etc.

Throughout the paper we will assume that, in equilibrium, the variables of the model have a law of motion that can be written as a time invariant function of a finite set of state variables, and we will assume that the researcher is interested in analyzing the stationary distribution of the model. The procedure can be adapted to cases where the law of motion depends on time (for example, Marshall [1988] solves by PEA a model with a time varying process for forcing variables), cases where the researcher is interested in the analysis away from the steady state (for example, Marcet and Marimon [1991] study by PEA the growth path from a low level to the steady state level of capital stock) and where it is not known if there is a finite number of state variables (for example, models with private information). Also, since we are interested in the steady state distribution, all the mathematical expectations in this paper are taken with respect to this distribution unless otherwise specified.

It is easy to write down systems like (1.1) for many interesting models, but it is much harder to write down the equilibrium law of motion of z_t. We think of (1.1) as a complete description of the economy; in other words, we have to include enough equations in this system so that only the solution $\{z_t\}$ that we are seeking satisfies

the above system of equations. In some models, there are side conditions that have to be satisfied in order for the system (1.1) to be sufficient: these side conditions may involve second order conditions, Kuhn & Tucker conditions, non-negativity constraints of some variables etc., and they can be considered part of the system g.

Notice that, by characterizing the model with a system like this one we can solve models without having to cast them in a planner's problem: problems with distorting taxes, monetary models, models with inequality constraints, private information etc. can be cast into (1.1) quite easily, but only some of them are the solution to a well known planner's problem

We now show some examples of models that fit (1.1):

Example 1.1 (Lucas [1978] asset pricing model)

A representative consumer chooses optimal paths of consumption of a perishable consumption good; we denote his consumption by c_t. This consumption good is produced by a tree that yields d_t units of the consumption good every period. Shares of ownership to the tree can be traded in a competitive market. The agent receives exogenous labor income w_t. The variables $\{w_t, d_t\}$ follow an exogenously given Markov process of order m, and they may have a continuous distribution.

The representative agent solves

$$(1.2) \qquad \max_{\{c_t, A_t\}} E_0 \left[\sum_{t=0}^{\infty} \delta^t u(c_t) \right]$$

$$\text{subject to} \quad . \quad c_t + p_t A_t \leq (p_t + d_t) A_{t-1} + w_t,$$

$$(c_t, A_t) \text{ measurable with respect to } \Omega_t.$$

where A_t denotes shares of ownership of the tree bought at time t, p_t is the price of each share in units of consumption good and Ω_t is the information set at t, which includes observations on all variables up to time t. The representative agent behaves competitively, taking prices as given.

Finally, we follow Lucas and introduce a shortselling contraint $A_t \geq K$, for some $K < 1$. Introducing this constraint has no effect on equilibrium, but its absence would make the maximization problem ill defined.

There is a net supply of one share, so that equilibrium in the

market for shares guarantees that

(1.3) $A_t = 1$ for all t .

The consumption good is traded competitively in spot markets, so that in equilibrium

(1.4) $c_t = d_t + w_t$.

The Euler equation for the maximization problem can be written as:

(1.5) $$u'(c_t)\, p_t = \delta\, E\, [\, u'(c_{t+1})\, p_{t+1} \mid \Omega_t\,] + \delta\, E\, [\, u'(c_{t+1})\, d_{t+1} \mid \Omega_t\,]$$

Let us map this model into (1.1). The vector z_t is given by

$$z_t = (p_t,\ c_t,\ A_t,\ d_t,\ w_t,\ d_{t-1},\ w_{t-1},\ \ldots,\ d_{t-m-1},\ w_{t-m-1})' ,$$

where we have included the endogenous variables and the relevant lags of the exogenous variables (recall that $\{w_t,\ d_t\}$ are Markov of order m), and u_t will be the innovations in $\{w_t,\ d_t\}$.

In this model, the system corresponding to (1.1) is given by (1.3), (1.4) and (1.5), plus the law of motion for $(d_t,\ w_t)$. Note that, by the spot-market version of Walras' law the budget constraint will then be satisfied; hence, this equation can be ignored.

Example 1.2 (Simple Stochastic Growth Model)

Let us consider the following generalization of the model by Brock and Mirman [1972]

$$\max E_0 \sum_{t=0}^{\infty} \delta^t \left[c_t^{1-\gamma} \right] / (1-\gamma)$$

(1.6) $$\text{s.t.} \quad c_t + k_t - \mu k_{t-1} = \theta_t k_{t-1}^{\alpha} \quad ,$$

where c_t denotes consumption, k_t is the capital stock and the productivity shock θ_t obeys the following process:

(1.7) $$\log(\theta_t) = \rho \log(\theta_{t-1}) + \varepsilon_t \quad ,$$

where ε_t is i.i.d., distributed $N(0,\sigma_\varepsilon^2)$, and $|\rho|<1$.[8]

The first order condition for optimality is

(1.8) $$c_t^{-\gamma} = \delta E_t\left[c_{t+1}^{-\gamma} (\theta_{t+1} \alpha k_t^{\alpha-1} + \mu) \right] \quad .$$

In this model, the set of variables in the economy is $z_t \equiv (c_t, k_t , \theta_t)$. The system g in equation (1.1) is given by the resource constraint (1.6), the law of motion of the productivity shock (1.7) and the Euler equation (1.8), and u_t is the innovation to the productivity shock ε_t.

One issue that may come up is if this system of equations is sufficient to characterize the solution, or if we need other side conditions; in particular one might be worried about satisfying the transversality condition. However, we know from dynamic programming that this model has a unique solution, and that this is uniquely given by the FOC and the requirement that the policy function is time invariant. Therefore, as long as we make sure that our solution follows a time invariant law of motion we know that the FOC is sufficient for optimality.

---

[8] This is also the model solved by the NBER Non-Linear Rational Expectations Modelling Group. The results for this and other solution methods are reported in the January 1990 issue of the Journal of Business Economics and Statistics. Den Haan and Marcet [1990] discuss the solution to this model by PEA in greater detail in that issue.

Example 1.3 (Asset Pricing Model with Habit Persistence)

Let us consider the same model as in Example 1.1, with the simplification that $w_t=0$ and that d_t is Markov of order one, but with the complication that the instantaneous utility function depends on lagged consumption. Now the agent's utility function is given by

$$E_0 \left[\sum_{t=0}^{\infty} \delta^t \, u(c_t, c_{t-1}) \right] .$$

This model has received considerable attention recently, among other authors, by Heaton [1990], Constantinides [1990] and Singleton [1991].

The conditions for equilibrium are now the same as in example 1.1 except that the Euler equation becomes

$$(1.9) \qquad \left[u_1(c_t, c_{t-1}) + \delta \, E \left[u_2(c_{t+1}, c_t) \mid \Omega_t \right] \right] p_t =$$

$$\delta \, E \left[\left[u_1(c_{t+1}, c_t) + \delta \, u_2(c_{t+2}, c_{t+1}) \right] (p_{t+1} + d_{t+1}) \mid \Omega_t \right] ,$$

where $u_i(.)$ represents the partial derivative of u with respect to the i-th term.

For this model

$$z_t \equiv (p_t, \, c_t, \, c_{t-1}, \, A_t, \, d_t)$$

and the system (1.1) is now given by (1.4), (1.5), (1.9) and the law of motion of d_t.

Example 1.4 (Simple Growth Model with Lower Bounds on Investment)

This example is meant to show that inequality constraints are easily handled by the PEA framework.

The simple growth model of example 1.2 is sometimes written with

inequality constraints; in particular, investment is often assumed to be non-negative. In this case we have to add the restriction

$$i_t \geq 0 \tag{1.10}$$

to the optimal problem in Example 1.2. With this restriction the first order condition (1.8) is replaced by the Kuhn & Tucker conditions:

$$i_t \left[c_t^{-\gamma} - \delta \, E_t\left[c_{t+1}^{-\gamma} \left(\theta_{t+1} \, \alpha \, k_t^{\alpha-1} + \mu \right) \right] \right] = 0 \quad , \tag{1.11}$$

plus the inequalities

$$c_t^{-\gamma} \geq \delta \, E_t\left[c_{t+1}^{-\gamma} \left(\theta_{t+1} \, \alpha \, k_t^{\alpha-1} + \mu \right) \right] \tag{1.12}$$

and (1.10).

In this case the vector z_t is the same as in example 1.2, but (1.11) and the inequalities replace (1.8) in that example. Hence, the model with inequality constraints can be put into the framework of (1.1), since the first order condition (1.12) can be imposed as part of the function g. Marcet and Singleton [1991] discuss in detail how to solve models with inequality constraints with PEA.

For a given model, there are many alternative systems (1.1). In particular, by letting current variables go in or out of the conditional expectation it is possible to represent a given model with different functions g and ϕ; for example, in Example 1.1 the Euler equation is often written as

$$u'(c_t) = \delta \, E \left[u'(c_{t+1}) \, (p_{t+1} + d_{t+1}) / p_t \mid \Omega_t \right] \quad ; \tag{1.13}$$

We will see below that this parameterization would not be useful for our approach, since it would make it impossible to obtain simulations for stock prices. More generally, the function g has to be written in a way so that given z_{t-1} and u_t, the system (1.1) provides a unique solution for z_t. In principle it is possible to find this solution since we have n equations for n unknowns.

Assume, for simplicity, that there is a finite vector of state variables x_t; then there is a function $f(.)$ such that

$$E\ [\ \phi(z_{t+1})\ |\ \Omega_t] \equiv f(x_t)\ .$$

If we knew the function f then we could substitute it in the first order conditions and (1.1) would become

$$(1.14) \qquad g(\ f(x_t),\ z_t,\ z_{t-1},\ u_t) = 0 \qquad \text{for all } t\ .$$

Clearly, (1.14) provides a solution for z_t in terms of z_{t-1} and the shocks, and this solution gives us the equilibrium law of motion of z_t. Hence, if we knew the function f, we could obtain simulations solving for z_t in (1.14) at every period. Unfortunately, the function f depends on the equilibrium process z_t ; if we knew the equilibrium law of motion for z_t we could probably derive f, but this law of motion is precisely what we are trying to find. This is the reason that solving models with rational expectations is complicated: the law of motion depends on the expectations, but we can not find the expectations until we have the law of motion.

The approach we propose avoids this vicious circle by starting with a parameterization of the conditional expectations in (1.1) as a fixed function $\psi(\beta,x_t)$, where β is a set of parameters, and x_t is a sufficient vector of variables in the information set; choosing x_t, ψ and β, so that $\psi(\beta,x_t)$ is as close as possible to $E\ [\ \phi(z_{t+1})\ |\ \Omega_t\ ]$. Substituting this approximated conditional expectation we have

$$(1.15) \qquad g(\ \psi(\beta,x_t(\beta)),\ z_t(\beta),\ z_{t-1}(\beta),\ u_t\ ) = 0 \qquad \text{for all } t$$

For a fixed choice of ψ, β and x_t it is easy to obtain a solution for $\{z_t(\beta)\}$ from this equation and, if the parameterized expectation is indeed close to f, then $\{z_t(\beta)\}$ is close to the equilibrium process for z_t. Of course, for an arbitrary selection of ψ, β, and x_t this solution will have little to do with the rational expectations equlibrium. The next paragraphs discuss how these selections should be made in order to have a good approximation. Notice that we make explicit the dependence of $\{z_t(\beta)\}$ on β, and we leave the dependence on x_t and ψ implicit.

Choice of x_t

The variables in x_t have to be contained in the information set of the agents and they should be rich enough so that, in equilibrium, $E(\phi(z_{t+1}) \mid x_t)$ is equal to the expectation in (1.1). In most models, we know enough about their solutions to make an informed guess about x_t; for example, in models where z_t can be obtained by solving a dynamic programming problem, x_t should be the vector of state variables.

Clearly, there are many choices of variables that satisfy the above requirement. For example, in the growth model of example 1.2 we could choose either $[k_{t-1},\theta_t]$ or $[k_t,\theta_t]$; since both of these choices form a sufficient set of state variables. Another consideration in choosing x_t is that, depending on the choice of x_t we can make to facilitate solving for $z_t(\beta)$ in (1.15). In Example 1.2, using the beginning of period capital stock we can solve for capital recursively, but using k_t would involve solving a non-linear system for finding k_t at every period. It is worth noting that the criterion described in this paragraph may conflict with the criterion in the previous paragraph.

Choice of ψ

The function ψ will be selected in this paper to belong to a class that can approximate arbitrarily well any function. Since we are trying to approximate conditional expectations, it is enough to approximate this function in the L^2 sense. A natural choice for this class of functions is a polynomial; but, in principle, other classes like splines or neural networks could be used. Again, it is important to choose ψ so that we can solve for $z_t(\beta)$ in (1.15).

If we decide to use a polynomial, the idea is to select a polynomial of degree high enough so that the function $\psi(\beta, \cdot)$ is close to the conditional expectation $f(.)$. One obvious way to check if the degree of the polynomial is high enough is to compare the solutions at different degrees of the polynomial; other ways of checking for accuracy are described in den Haan and Marcet [1989]. If the researcher were using splines, the researcher would have to use more refined partitions of the state space in order to obtain arbitrary accuracy.

We can discuss the choice of x_t and ψ in the context of the previous examples:

Example 1.1

One candidate for a state variable would be the stock holdings A_t, because this is a state variable in the maximization problem of the agent. But we know that, in equilibrium, A_t is a constant so that it carries no information on future variables; therefore, we do not have to include it in x_t. Also, only m lags help predict future values of the exogenous variables. Hence, we choose x_t equal to the sufficient set of state variables

$$x_t = (d_t, w_t, d_{t-1}, w_{t-1}, \ldots, d_{t-m-1}, w_{t-m-1})'$$

Notice that there are two expectations that have to be approximated[9]. One possible choice for ψ in each expectation is a regular polynomial.

Example 1.2

Per our previous discussion, we could choose $x_t = (k_{t-1}, \theta_t)$.

As for the choice of functions, we note that the left hand side of (1.8) is strictly positive, as well as the functions inside the conditional expectation. In this case it is advisable to choose a function ψ that can only take positive values; if we used a regular polynomial, this could take a negative value and make it impossible to solve for $c_t(\beta)$. Also, forcing ψ to be positive captures a feature of the true conditional expectation which, we know, can not take negative values. One alternative is to use the exponentiated polynomials $\exp(P_n(x_t))$, which can only take positive values and can approximate arbitrarily well the true conditional expectation.

Example 1.3

This is an interesting case, because even though the set of state variables of the model is (c_{t-1}, d_t), the conditional expectations

---

[9]These two expectations could be grouped into one. Usually it is advisable to reduce the number of expectations when doing numerical work, since this reduces the number of parameters that we have to iterate over. We separate the two expectations in (1.5) because this facilitates the analysis of S and β_f later in this section.

involve values of future and present values of the dividend process and, since d_t follows a Markov process of order one, these expectations depend only on the current value of the dividend. Hence, we can choose

$$x_t = d_t \quad .$$

Example 1.4

Here the state variables are as in Example 1.2. In this case, there is no need to force ψ to be positive; a negative value of the expectation simply means that, for that period, the solution is $i_t=0$, so that the equation (1.11) and the inequality (1.12) are both satisfied.

Choice of β

Once we choose a functional form we have to decide the coefficients β in this function. For example, if we choose a polynomial, β represents the coefficients in the different terms of the polynomial; if we choose a spline, β will contain the partition points as well as the coefficients in each interval.

The idea is to choose β so that, if we find the solution with this parameter, this is the parameter that best predicts $\phi(z_{t+1}(\beta))$. This is justified because, if $f(x_t)$ represents the true conditional expectation under rational expectations, then this function minimizes the mean square error when we try to predict $\phi(z_{t+1})$ with a function of x_t; more formally, it satisfies

$$f = \underset{h \in F}{\text{argmin}}\ E \,||\, h(x_t) - \phi(z_{t+1}) \,||^2 \quad ,$$

where $F \equiv \{\ h:R^n \rightarrow R^m\ \}$. Given the solution $\{z_t(\beta)\}$, we can ask the following question: what is the parameter value that minimizes the mean square error when we predict $\phi(z_{t+1}(\beta))$, if we restrict ourselves to functions of the form $\psi(\ .\ , x_t(\beta))$?. This is given by the following mapping:

$$S(\beta) = \underset{\beta^*}{\text{argmin}}\ E \,||\, \psi(\beta^*, x_t(\beta)) - \phi(z_{t+1}(\beta)) \,||^2 \quad . \tag{1.16}$$

Here, S maps R^{ν} into itself, and ν is the number of coefficients in β. The mapping S can be interpreted as the mapping from perceived to optimal expectations, in the sense that if agents perceive $\psi(\beta,x_t)$ to be the best way to forecast $\phi(z_{t+1})$ with information at t, in fact, the best forecast that can be made if we are restricted to forecasting with the function ψ, is given by $\psi(S(\beta),x_t)$.

We then choose $\beta=\beta_f=S(\beta_f)$. By definition of the mapping S, the fixed point is such that, if agents use β_f in their forecasts, then the series that satisfies the first order conditions will confirm that this is the best forecast that agents could possibly use within functions of the form $\psi(\cdot,x_t)$. Therefore, our numerical solution is given by $\{z_t(\beta_f)\}$.

It is beyond the scope of this paper to prove existence or uniqueness of the fixed point. In particular, one could be suspicious about uniqueness of the fixed point, because we know that if $\{g_n\}$ is a sequence of polynomials such that $g_n \longrightarrow h$ as $n\longrightarrow\infty$, even if h has a unique fixed point, g_n will have, in general, many fixed points. However, the fact that ψ could be polynomial which will have stronger oscillations for high degrees does not mean that S displays similar behavior when we increase the degree of the polynomial. This is illustrated in the following example, where we calculate S and β_f analytically for the model in Example 1.1; in particular, we will show that S is linear and it has a unique fixed point for any degree of the polynomial.

Example 1.1 Stability of S, Existence and Uniqueness of β_f

The key element of this model that facilitates the analysis is that x_t consists only of exogenous variables. Equations (1.3) and (1.4) immediately give us the law of motion for c_t and A_t, so that the only variable that we have to solve for is p_t.

The parameterized expectation will take values in R^2, since two conditional expectations are involved in (1.5). We choose each element of ψ to be a regular polynomial of degree ν. Then we can express ψ as

$$\psi(\beta,x_t) = [\ \beta_1,\ \beta_2\ ]' \cdot f(x_t),$$

where $f(x_t)$ is a vector containing all terms of the form

$(x_{1t})^{q} \cdot (x_{jt})^{q'}$, for integers $q, q' \geq 0$ such that $q+q' \leq \nu$, and for x_{1t}, $x_{jt} \in x_t$. Here, β_1 and $f(x_t)$ are column vectors of the same length.

Now we substitute the conditional expectations in (1.5) with the parameterized expectations, to obtain

$$u'(c_t)\, p_t(\beta) = \delta\, (\beta_1 + \beta_2)'\, f(x_t) \quad . \tag{1.17}$$

This is the system (1.15) in this model.

Now, observe that

$$\hat{E}\, [\, u'(c_{t+1})\, p_{t+1}(\beta) \mid f(x_t)\,] = S_1(\beta)'\, f(x_t)$$

$$\hat{E}\, [\, u'(c_{t+1})\, d_{t+1} \mid f(x_t)\,] = S_2(\beta)'\, f(x_t),$$

where $\hat{E}(X|Y)$ denotes the linear projection of X on Y. This equality follows from the definition of S, the definition of a linear projection and because ψ is linear in $f(x_t)$.

Since c_t, d_t, and $f(x_t)$ are all exogenous variables, $S_2 \equiv \gamma$ for a vector γ independent of β. Finally, we find a formula for S by observing that if p_t is given by (1.17), then

$$\hat{E}\, [\, u'(c_{t+1}) p_{t+1}(\beta) \mid f(x_t)] = \hat{E}\, [\, \delta(\beta_1 + \beta_2)' \cdot f(x_{t+1}) \mid f(x_t)] =$$

$$\delta(\beta_1 + \beta_2)'\, \Gamma\, f(x_t) \equiv S_1(\beta)'\, f(x_t),$$

where the square matrix Γ satisfies $\hat{E}(f(x_{t+1}) \mid f(x_t)) = \Gamma\, f(x_t)$, and where

$$S_1(\beta) = \delta\, (\beta_1 + \beta_2)'\, \Gamma \tag{1.18}$$

$$S_2(\beta) = \gamma$$

Also, it is easy to check that the fixed point β_f satisfies

$$\beta_{1f} = (\, I - \delta\, \Gamma'\,)^{-1}\, \Gamma'\gamma$$

(1.19)

$$\beta_{2f} = \gamma \quad .$$

These are analytic formulas for S and β_f up to knowledge of γ and Γ. It is in this sense that we say that in this model S and β_f can be solved analytically. The parameters γ and Γ will usually have to be calculated numerically by running the appropriate OLS regressions. This formula illustrates the point that S is well behaved: for any degree of the polynomial ν, S is a linear mapping with a unique fixed point. Also, these formulas will help in establishing stability under learning and convergence of iterative algorithms for this model.

In the next section we will how $S(\beta)$ can be calculated by running non-linear regressions of $\phi(z_{t+1})$ on $\psi(.,x_t)$; taking this into account, we can finally describe the algorithm as a sequence of steps:

Step 1: Given the system (1.1), choose x_t and the class of functions ψ appropriately; substitute the expectations in the system (1.1) by the parameterized expectations $\psi(\beta,x_t)$ in order to obtain (1.15).

Step 2: For a given parameter β obtain a simulation $\{z_t(\beta)\}_{t=1}^{T}$ for large T, using (1.15).

Step 3: Run a non-linear regression of $\phi(z_{t+1})$ on $\psi(.,x_t)$; the result is (an approximation to) $S(\beta)$.

Step 4: Find the fixed point $\beta_f = S(\beta_f)$.

The Solution as an Approximation and Comparison with other Algorithms

The intuition of why this is a good approximation to the rational expectations equilibrium is clear: we choose the polynomial that satisfies the crucial property of the conditional expectation, namely, that it minimizes the mean square prediction error. A rigorous proof that this procedure in fact approximates arbitrarily well the true solution is difficult; one can not invoke directly the Stone-Weierstrass theorem because the variables that are being predicted themselves depend on the parameters of the polynomial; also, there is the issue of the non-linear regressions being only an approximation to $S(\beta)$, and the issue of having initial conditions of the state variables that are not drawn from the steady state

distribution. A proof of convergence when x_t contains only exogenous variables (as in Examples 1.1 and 1.3) can be found in Marshall [1988]; a more general proof is the object of current research.

Independently from the approximation properties of the solution when the polynomial is arbitrarily large, the solution for a given polynomial degree can be interpreted as the equilibrium when agents are restricted to using a certain functional form in forming their conditional expectations. So, if the researcher has to use a low degree polynomial due to computational limitations, at least he can be confident that the solution he obtains is interpretable as a well defined equilibrium concept.

One important advantage of this algorithm is that it works by doing endogenous oversampling. One common problem of algorithms that use grids in the state space is that, since the researcher does not know beforehand what values of the state variables will happen more frequently, he may impose a very large grid, but only a few points in this grid are important. This problem could be overcome by spending some computer time updating the grid, trying to adapt to the solution, and giving more importance to those points that happen more frequently. PEA does this procedure automatically, because the polynomial is fitted on the domain of the variables that happen often in the long run simulation. The problem with endogenous oversampling may be that, since the computer has to search what area of the state variables is rellevant, the algorithm may end up looking in areas that are irrelevant; more precisely, the long run simulations may be explossive. In other words, endogenous oversampling may leave too much work to the computer. This can be solved in part by the homotopy approach discussed in the next section, and by changing the long run simulations by repeated short run simulations as in Marshall [1988] and Marcet and Marimon [1991].

Also, PEA can be thought of as using a Monte-Carlo method to calculate the integrals (more precisely, the expectations) in the model, since the long run simulations are used to determine $S(\beta)$ which is, in fact, determined by expectations. Monte-Carlo integration is useful, particularly in models with several exogenous random shocks, where quadrature may be too computationally intensive. Notice that the endogenous oversampling and the Monte-Carlo integration are done jointly in Step 2, when the long run simulation is calculated.

Similar ideas for finding solutions to dynamic models have been presented in the literature. Some of the ideas on how to set up this problem and how to find the fixed point β_f were present in Marcet and Sargent [1989a, 1989b]. Also, they showed that stability of the S-mapping determined convergence of least squares learning schemes in linear rational expectations models.

The method of Townsend [1983] can be thought of as selecting ψ to

be linear, and performing particular kinds of iterations on the mapping S in order to find β_f. The setup in this paper separates the problem of specifying ψ and x_t from the problem of finding the fixed point β, allowing for more flexibility in the class of models that can be addressed.

The backwards solution method of Sims-Novales-Ingram can also be interpreted in light of the above framework. This method involves guessing about the true process for the endogenous variables, and using this guess to substitute the expectation in the Euler equation by a convenient expression; then we can solve for the exogenous variables. The "guess" about the process for the endogenous variables can be thought of as guessing for ψ, x_t and β_f. One of the main ideas of PEA is that we should substitute the conditional expectations in the Euler equations by a 'convenient' expression that makes it possible to solve for the endogenous series, an idea that has been taken from the backwards solution procedure.

The Euler equation methods of Coleman [1989], Bizer and Judd [1989] and Judd [1990], parameterize the law of motion and iterate on this law of motion until the Euler equation is satisfied at a grid imposed on the space of state variables; in these papers,[10] the expectations in the Euler equations are calculated by quadrature. In general, since PEA does Monte-Carlo integration, one would expect it to be faster in the presence of several exogenous random shocks. Also, it has the advantages (and disadvantages) of endogenous oversampling. Finally, parameterizing the law of motion or the expectation may be a matter of taste; some advantages of parameterizing expectations are that the solution is then interpretable as the limit of a well specified learning process and that the number of state variables is sometimes smaller, since we are only worried about their predictive power (this is the case in example 1.3 above and in Marcet and Marimon [1991]). Also, in models with private information we know that additional lags, or additional functions of the observed processes, are relevant if they have predictive power, so that this gives us a guideline as to what terms should be introduced in the expectation to obtain a better approximation.

---

[10] Judd shows how these procedures and PEA can be interpreted in terms of a more general framework. He also points some similarities with numerical procedures used in physics and engineering.

2. Numerical Solution to the Fixed Point of S.

In this section we discuss how to find the fixed point of S for a given choice of ψ and x_t. The first problem is how to find $S(\beta)$. Notice that we can simulate a long series of $z_t(\beta)$ from equation (1.15). Also, $S(\beta)$ satisfies

$$E\left[\left[\phi_i(z_{t+1}(\beta)) - \psi_i(S_i(\beta),x_t(\beta) \right] \left. \frac{\partial \psi_i(\beta^*, x_t(\beta))}{\partial \beta^*} \right|_{\beta^*=S_i(\beta)} \right] = 0.$$

$$\text{for } i = 1,\ldots,m$$

These are just the first order conditions corresponding to the minimization problem (1.16).

Hence, given a series $\{z_t(\beta)\}_{t=1}^{T}$, we can run a non-linear least squares regression (NLR) of $\phi(z_{t+1}(\beta))$ on the function $\psi(\cdot,x_t(\beta))$. Since the non-linear regression can be interpreted as the GMM estimator with orthogonality conditions given above, we know that as T goes to infinity the result of the non-linear regression converges to $S(\beta)$. Thus, we can calculate $S(\beta)$ with arbitrary accuracy by using T arbitrarily high. In practice, one can check if a given T is large enough by comparing the solution with several realizations of the shocks.

Compared with algorithms that use quadrature to evaluate expectations, calculating S with the above procedure is related to using Monte-Carlo integration to evaluate the expectations in the first order conditions for the minimization problem. Using Monte-Carlo methods is a very efficient way to calculate expectations in a model with several exogenous random variables, which is the usual case except in extremely simplified models. While the cost of performing quadrature integration grows very rapidly when the number of shocks increases, the calculations involved in the regression are unaffected by the number of underlying shocks in the economy.

The non-linear regression does become more expensive if we increase the number of state variables. However, unlike in grid-based methods, the cost does not grow exponentially: for a polynomial of degree ν, the number of coefficients in the regression increases at a rate less than n^{ν}, where n is the number of variables in x_t. The next paragraph argues that, in practice, we usually have to calculate much fewer terms than that.

In adding higher order terms of the polynomial, it turns out that many of these terms do not have any predictive power over the terms that already existed. For example, in the growth model of example 1.2, in going from a first degree to a second degree polynomial, it turns out that the cross term $k_{t-1}\theta_t$ does not have any predictive power when the other second order terms are introduced, so this term does not have to be introduced and there are fewer parameters involved in β_f. Whether these terms have to be introduced or not can be checked before they are actually introduced, as discussed in den Haan and Marcet [1989]. In this way it is possible to handle relatively large state vectors.

The next step consists of finding an iterative scheme that will allow us to find the fixed point of S.

An Iterative Scheme to Find β_f

We will discuss algorithms that converge if the least squares learning mechanism is stable. Consider the differential equation

$$\dot{\beta} = S(\beta) - \beta. \tag{2.1}$$

Along solutions to this differential equation, β is adjusted infinitesimally towards $S(\beta)$. Clearly, the only equilibrium of (2.1) is at β_f.

Marcet and Sargent [1989a, b] show that in the linear case and under various assumptions concerning the variables in the information set, least squares learning schemes are locally stable if, and only if, the above differential equation is locally stable. Some of these results have been extended to the non-linear case by Bansal [1988]. We will now prove that for the model in Example 1.1, (2.1) is stable and the least squares learning mechanism is stable.

Since S is a linear mapping, (2.1) is a linear differential equation, and all we need to show in order to prove global stability of (2.1) is that the eigenvalues of the derivative of S are all less than one. But the eigenvalues of S are all equal to $\delta \cdot \xi_1$, where ξ_1 is some eigenvalue of Γ. We already argued at the end of the last section that $|\xi_1| \leq 1$; since $0<\delta<1$, we have that (2.1) is stable, and the rational expectations equilibrium in Lucas' asset pricing model is stable under least squares learning. Since the variables used in predicting the future are exogenous, using an argument similar to Corollary 2 of Marcet and Sargent [1989a] we can prove that the least

squares learning mechanism converges.[11]

The relationship between (2.1) and stability of learning schemes justifies using algorithms based on this differential equation. More precisely, we will perform iterations on β of the form

$$(2.2) \qquad \beta_{\tau+1} = (1-\lambda)\,\beta_{\tau} + \lambda\, S(\beta_{\tau}), \qquad \text{for some } 0<\lambda \quad .$$

If these iterations converge, they converge to β_f. If they do not converge even for very small λ, it can only happen because the rational expectations equilbirium is locally unstable under least squares learning.

Obviously, we could also use some gradient algorithm for finding solutions to non-linear systems of equations. These algorithms follow more direct routes to the solution but, since the gradient has to be calculated, the usual algorithms will often be too computationally intensive for the problems we consider. The algorithm (2.2) follows a zig-zag path to the solution but, since it is very easy to update, it is usually faster. Finally, and no less important, algorithm (2.2) is extremely easy to program.

A Fast Stochastic Algorithm.

Another way of calculating β_f is using an algorithm from the stochastic approximation literature. The model we simulate is given by the following equations:

$$(2.3) \qquad \beta_{t+1} = \beta_t + \alpha_t\, R_t^{-1} \left[D\psi_t' [\, \phi(z_{t+1}) - \psi(\beta_t, x_t) \,] \right]$$

$$(2.4) \qquad R_{t+1} = R_t + (1/t) \left[D\psi_t \cdot D\psi_t' - R_t \right] \equiv (1/t) \sum_{i=0}^{t} \psi_t \cdot \psi_t'$$

---

[11] Strictly speaking, the model of section 2 does not fit the Marcet and Sargent framework because $f(x_t)$ is a non-linear function of x_t, but it does fit the framework of Bansal [1988].

(2.5) $\quad g(\ \psi(\beta_t, x_t),\ z_t,\ z_{t-1},\ u_t\ ) = 0$

$$\text{where } D\psi_t = \frac{\partial\ \psi(\beta_t\ , x_t)}{\partial\ \beta}\ , \qquad \text{for all t.}$$

This can be interpreted as a model of learning in a non-linear environment. If we choose $\alpha_t = 1/t$, equations (2.3) and (2.4) are a form of the recursive non-linear least squares estimator given the expectation function ψ. The coefficients β_t are updated every period using a new observation, and this coefficient is used in the forecasts at t, in equation (2.5). Schemes like this have been used in the adaptive control literature, and they are discussed thoroughly in Ljung and Söderström [1983].

It is well known from the literature of stochastic approximation (see for example Ljung and Söderström) that algorithms of the type (2.3)-(2.5) only converge to points β^* satisfying

$$E\ \Big[\ [\phi(z_{t+1}(\beta^*)) - \psi(\beta^*, x_t(\beta^*))\ ]\ D\psi_t(\beta^*)\ \Big] = 0 \quad [12].$$

Then, by definition, $\beta_f = \beta^*$ and this recursive scheme can only converge to β_f. This recursive algorithm has the property that it converges to the fixed point that we are seeking 'automatically', without having to perform the regressions and the iterations on β_τ separately.

This is why the latter algorithm can be less computationally intensive than performing the iterations in (2.2). There, every time that $S(\beta_\tau)$ is evaluated, we have to go through the simulated data several times in order to perform the non-linear regression, while in the above scheme we only use one data point at each period. The advantadge of performing iterations in (2.2) is that they are non-stochastic, so that it is easier to determine if we are at a fixed

---

[12]To obtain the left hand side of this equation, we take the expectation of the element multiplying α_t, where the expectation is taken with respect to the stationaty distribution of $z_t(\beta)$. See Marcet and Sargent (1989a,b) for another application of this recursive approach to least squares learning models.

point, while the stochastic nature of the algorithm (2.3)-(2.5) makes it harder to control and to determine if we have arrived at a fixed point.

In practice, we find it is best to start calculating solutions to a given model with algorithm (2.2); once we feel comfortable with the model and the nature of its solution, we can move to the recursive algorithm of (2.3)-(2.5) to do the more computationally itensive tasks.

Also, it turns out that using $\alpha_t=(1/t)$ is not very efficient; the reason is that these weights go to zero too fast; while this is what warrants convergence of the algorithm it also makes convergence very slow, because too little importance is given to the new observations. In the engineering literature it is common to set $\alpha_t = \varepsilon$, where ε is a fixed small number; this allows "tracking" systems that may experience large changes in the structure generating the data. Of course, the problem is that unless α_t goes to zero there is no chance for β_t to converge; the solution is to set α_t constant in a few periods and then let it go to zero. In the next section we describe how to do that for a particular growth model and how this can be used to track changes along a homotopy.

These two algorithms, will converge whenever the model is stable under learning. In principle, it would seem desirable that our models would have some type of stability properties under learning schemes; furthermore, recent findings in the learning literature indicate that, even though it is possible to find models where learning is unstable, these are more the exception rather than the rule.

Even though local stability of these algorithms is often warranted, global stability is harder to check. More importantly, it is possible that in the course of the iterations of the form (2.2) β_τ will enter a region where z_t has explosive paths, making $S(\beta_\tau)$ ill-defined and causing our algorithm to break down; this can happen because the non-linear difference equation implied by (1.15) may, be explossive for arbitrary choices of β, even though it will be stable near the fixed point if this is a good approximation to the true conditional expectation.

Using the Homotopy Approach.[13]

The rest of this section addresses the problem just described using the homotopy approach. In effect, this is just a way of finding

---

[13] Guido Tabellini pointed to me the usefulness of this literature.

good initial conditions for the iterations on β_f and avoiding explossive solutions during the iterations. Our experience indicates that, for most models where endogenous variables enter in x_t, it is necessary to use this approach.

Let us briefly discuss the homotopy approach in general. We follow in part the exposition of Garcia and Zangwill [1981]. Assume that we want to solve the following system of equations:

$$F(x,\alpha) = 0$$

for $F:R^n x R^m \to R^n$, for some α^*. Denote the solution x^*, so that $F(x^*,\alpha^*) = 0$. Assume that we know that for some parameter α_0 the solution is x_0, so that $F(x_0, \alpha_0) = 0$. The idea is to move gradually from the known solution x_0, to the solution that we are interested in x^*.

Define a homotopy $H:R^n x[0,1] \to R^n$ as a function that satisfies

$$H(x_0,0) = 0 \quad \text{and} \quad H(x^*,1) = 0.$$

A path $x(\tau)$ satisfies $H(x(\tau), \tau) = 0$ for all $\tau \in [0,1]$. Under some continuity assumptions on H it is easy to guarantee that this path exists. Then we can use algorithms that necessitate "good" initial conditions and, letting τ go gradually from 0 to 1, we will follow this path to move gradually from the known solution $x(0) \equiv x_0$ to the solution we seek, $x(1) \equiv x^*$.

The idea of homotopy is very simple, but it is also extremely powerful. In many models it possible to go from a known solution to the solution we are seeking by using our knowledge of economic theory. If this fails, one can draw from a large literature in mathematics on how to construct homotopies.

In the simple growth model of example 1.2, the case when γ=1 and μ=0 is the one solved by Brock and Mirman [1972], so we can go from this known solution to the desired parameter values (usually, for yearly data, we would like to set μ=.9 and the risk aversion parameter should be higher than one). Since we know the solution of the Brock and Mirman model analytically, we know β_f for those parameters, and we use it as an initial condition for the homotopy.

More generally, we can create a homotopy if the solution we know and the solution we seek are both nested in a more general model, and they correspond to different parameter values for the same model. Formally, H can be defined as follows. It was implicit in the definition of S that this mapping, and therefore its fixed point,

depend on the values of the parameters of the model[14]. Letting α represent the parameter values of the model, and writing the dependence of S on α explicitely, we have the mapping $S(\alpha,\beta)$ with the fixed point $\beta_f(\alpha) = S(\alpha,\beta_f)$. Assume that we know the solution of the fixed point for a given set of parameters α_0 is β_{0f} and we want to calculate the solution for α^*; we can set up the homotopy as

$$H(\beta,\tau) = S(\alpha^*\tau + \alpha_0(1-\tau), \beta) - \beta ;$$

clearly, for $\tau=0$ the solution is β_{0f}, and for $\tau=1$ it is the solution we seek.

Generically, continuity of the homotopy guarantees existence of the path almost everywhere. If g and ϕ in equation (1.1) are continuous, S is continuous; then we know that the path will generically exist, and so will the intermediate fixed points. This assumes that the model is well defined at the intermediate point of the path.

For another example, consider the case where we just want to change the instantaneous utility function from *h* to another functional form *f*, and these functions are not a special case of each other. For example, *f* could be a CRRA utility function, and *h* could be the exponential utility function. Then, we could construct a homotopy by solving the model with the utility function: $(1-\tau)\ g + \tau\ f$. Clearly, by letting τ go from zero to one the solution moves to the solution we are seeking. It is worth noting that this function preserves the property of concavity of utility functions, so that the 'intermediate' solutions are well defined. This approach is related to the 'linear homotopy' described by Garcia and Zangwill.

Another useful homotopy is obtained by starting the solution by fixing the whole stochastic process z_t as an arbitrary linear stationary process, depending on extraneous shocks, and moving to the true solution. More precisely, we would obtain $z_t(\beta)$ by solving

(2.6) $$\tau\ g(\ \psi(\beta,x_t(\beta)),\ z_t(\beta),\ z_{t-1}(\beta),\ u_t\ ) +$$

---

[14] More precisely, the functions g and ϕ and the process for the exogenous variables depend on the parameter values.

$$(1-\tau)\ [\ z_t(\beta) - A\ z_{t-1}(\beta) - \eta_t\ ] = 0 \qquad \text{for all } t \quad ,$$

where η_t are some extraneous shocks, unrelated to u_t, and where A is a matrix with all eigenvalues less than one in modulus. Again, letting τ go from zero to one we will go from the trivial solution of the difference equation to the solution we seek.

3. An Asset Pricing Model with Endogenous Production.

Different versions of the asset pricing model with exogenous production have been widely explored in the literature. In their seminal paper, Mehra and Prescott [1985] (MP) argued that, for reasonable values of the risk aversion parameter, the equilibrium model described in Example 1.1 could not explain the equity premium observed in real data. They assumed that dividends equalled consumption and that it was a two-state Markov chain. Since then, a large number of papers has attempted to explain this puzzle.

The literature has focused on models with exogenous production and has tried to find more interesting preference structures or dividend processes[15]. While part of this focus may be due to technical limitations, it may also due to the belief that a model similar to Lucas' where production is made endogenous by explicitely modelling the capital accumulation process will only make it harder for the model to match the data, since it means increasing the number of restrictions imposed on the data. Since asset prices only depend on the dividend and consumption processes, introducing capital only makes it harder to match the data.[16]

Recently, there has been a renewed interest in modelling explicitely production in order to explain asset prices; some references are Rowenhorst [1989] and den Haan [1990b]. In this section we argue that modelling production explicitely may help explain some features of the data; in particular we will show that the model can produce a humped-shaped term structure of interest rates precisely in recessions, somehow agreeing with the observations of Fama [1984] on the empirical behavior of interest rates, while the model with exogenous production can not generate this observation. Also, we will study the risk premium when stock returns are made highly volatile. Finally, Schwert [1989] shows empirical evidence that the volatility of stock returns is higher in recessions, and Rowenhorst [1989] argues that the model with endogenous production is able to generate this observation; we will point out that, in fact, most equilibrium models display this behavior.

Another advantage of modelling production explicitely is that we have a well articulated economic structure that justifies the dividend and labor income process. This means that we can study the relationship between the stock market and real economic activity, like investment, productivity shocks, etc. We will use this feature of the model to create a model with very few parameters where the volatility

---

[15]See, for example, Riesz [1988], Backus, Gregory and Zin [1989] and Epstein and Zin [1989].

[16]See, for example, Mehra and Prescott [1985].

of stock returns matches that of the data, a feature not displayed by the Mehra and Prescott model, where consumption equals dividends.

Trying to construct a model of this kind that explains all aspects of the data would be beyond the scope of this paper; the reader is referred to Rowenhorst [1989] and den Haan [1990b] for a much more detailed analysis. Instead, we will show how, with very few parameters, we can generate rich patterns of asset prices. The model fails to explain the risk premium in the data, but the risk premium increases considerably compared with the one that would happen if consumption was equal to dividends, even with a standard deviation of stock returns that matches the data.

The rest of this section is organized as follows: first we discuss how to solve for the real variables of the problem, then we specify the model with securities and show how to price them, and finally we study the properties of the model.

Solving a Simple Growth Model

The model will be a version of Brock [1982], and will be fully specified in the next subsection. As Brock showed, the real part of the economy is described by the simple growth model of Example 1.2, so that the real variables, namely, consumption, investment and capital are given by the solution to that model. So we will first discuss how to solve for the real variables in Example 1.2.

The state variables are known to be (k_{t-1}, θ_t), which we take as our vector x_t; parameterizing the expectation in the right side of the Euler equation we obtain

$$c_t(\beta)^{-\gamma} = \delta \psi(\beta, k_{t-1}(\beta), \theta_t) \quad . \tag{3.1}$$

We choose as our functional form for the parameterized expectation

$$\psi(\beta, k_{t-1}, \theta_t) \equiv \exp [P_r(\log(k_{t-1}), \log(\theta_t))] , \tag{3.2}$$

where P_r is a polynomial of degree r. The main reason for this choice is that ψ takes only positive values: since ϕ is positive valued, the conditional expectation $f(.)$ has non-negative values and by imposing the same restriction on ψ we can presumably approximate f better.

This choice does satisfy the requirement that it can approximate the function f with arbitrary accuracy: letting

$$ff(x_1,x_2) \equiv f(\exp(x_1), \exp(x_2)),$$

we have that $ff(\log(k_{t-1}),\log(\theta_t))$ equals the conditional expectation, so it is enough to argue that the function $\exp(P_r)$ approximates ff. A polynomial P_r can approximate the function $\log(ff)$ arbitrarily well; hence, $\exp(P_r)$ can approximate ff arbitrarily.[17]

Another practical reason for this choice of ψ is that the right hand side of (3.2) is guaranteed to be positive, so that it is easy to solve for consumption.

Finally, we can set up a homotopy that goes from the Brock and Mirman solution with γ=1 and μ=0 to the solution we seek with lower depretiation

Let us go through the steps of the algorithm as described in section 1, using an exponentiated polynomial of degree one. Then (3.1) becomes

$$c_t(\beta)^{-\gamma} = \beta_1 \, k_{t-1}(\beta)^{\beta 2} \, \theta_t^{\beta 3} \tag{3.3}$$

so that (1.6), (1.7) and (3.3) form the system (1.15) in this model. This completes Step 1.

Obtaining solutions for consumption and capital for a given β is trivial: at time t the state variables are predetermined, so $c_t(\beta)$ is obtained directly from (3.3) and $k_t(\beta)$ from the production constraint; this completes Step 2.

Step 3 consists of running a non-linear regression trying to

---

[17]All of these approximations are in a compact set of the space of the state variables, so that we can apply the Stone-Weirstrass theorem. Strictly speaking there is no guarantee that our state variables will be on a compact set with probability one at all periods, but this is not a practical problem as long as we are solving a model with an ergodic distribution. If one wanted to be more strict, it would be possible to simulate a modified model with one additional constraint that prevents the capital stock from being higher than a very large upper bound; since inequality constraints are easy to handle with PEA this would be easy to simulate, while a model with a very large upper bound is essentially the same model as without the constraint and the solution should be very similar, as long as the capital stock has a stationary distribution.

predict

$$c_{t+1}(\beta)^{-\gamma} \left(\theta_{t+1} \alpha k_t(\beta)^{\alpha-1} + \mu \right) \quad \text{with} \quad \psi(\cdot, k_{t-1}(\beta), \theta_t) ,$$

and the result would be our approximation to $S(\beta)$[18]. Then it is straightforward to apply the first algorithm in section 2 to find the fixed point in Step 4 of the algorithm.

The following table summarizes calculations of the fixed point and computation times for the homotopy that goes from the solution we know (Brock and Mirman) to the solution to the solution we seek μ=1. [19]

μ	β_{1f}	β_{2f}	β_{3f}	time(sec.)
0.0	1.53	-0.33	-1.00	
0.1	1.49	-0.35	-0.97	10
0.2	1.45	-0.37	-0.94	8
0.3	1.42	-0.39	-0.91	10
0.4	1.40	-0.41	-0.87	8
0.5	1.39	-0.43	-0.83	12
0.6	1.40	-0.46	-0.78	8
0.7	1.44	-0.49	-0.72	8
0.8	1.53	-0.52	-0.65	18
0.9	1.74	-0.56	-0.55	25
1.0	2.47	-0.65	-0.36	65

To calculate the equilibrium with a higher polynomial degree we could use the solution for the first degree polynomial to run a regression of ϕ on a second degree polynomial, and use the result to start the iterations for the fixed point of the 2d degree exponential polynomial. Using this initial condition is better than starting the iterations at the previous solution with zeros in the second degree terms and it only takes one regression per expectation to find these

[18] It is not correct to run a linear regression on the logs, because the error term in the expectation is not additive.

[19] This table is taken from den Haan and Marcet [1990]. The remaining values of the parameters in the model are δ=.95, σ_ε=.1, λ=1. We used 2500 observations to run the non-linear regressions and four digits of accuracy to determine covergence of the iterations on β. The computer was a Compaq 386, 25 megahertz and a floating point processor.

initial conditions. The solution changes slightly when going to the second degree polynomial, indicating some relevant non-linearities in the problem; notice that the standard deviation of ε_t is quite large, which accounts for the importance of the non-linearities.

When we used the stochastic algorithm of Section 2, first we set α_t=.001 and we start at μ=0. Then we move the depretiation to μ=1 in steps of .1 and, because of the tracking characteristics of the algorithm, we get fairly close to the fixed point for μ=1. Finally, to make it actually converge, we set α_t= 1/t. All of this process takes about five to six times less than the total of the above computing time in the previous table. The drawback, is that we had to experiment considerably with the level of fixed α that would be stable but would converge fast enough, and we had to experiment with the time that we switch to the weights 1/t.

We also calculated the fixed point with the Gauss-Newton method using numerical derivatives. The fact that the derivatives had to be calculated numerically made this algorithm two to three times slower than the simple algorithm described in the previous table.

The emphasis of this paper is not on obtaining the maximum speed for the algorithm; there are many techniques that could be used to increase computational speed but, to keep things simple, we are concentrating on the simplest version of PEA as described in Sections 1 and 2. Nevertheless, and given that some researchers have questioned the speed of PEA, we are forced to say something about this not too exciting issue. First of all, we note that comparisons of speed should be made with extreme care; it is known to anybody who has ever used numerical algorithms that the speed reported in a paper can be increased dramatically by choosing appropriately some parameters in the calculations. For example, the above computation times would be cut by a factor of four if one used an accuracy criterion on the fixed point of .001 instead of .0001, and if one used half the number of periods in the simulation, while the solution would be virtually unchanged.

Also, the computation times are greatly reduced by choosing the initial conditions shrewdly. The above table does not choose initial conditions shrewdly; it is intended to demonstrate the homotopy idea of moving gradually from a known solution to the desired solution, and it chooses initial conditions in a terrible way if one is interested in speed. The above solution uses the initial conditions for the case where μ=0 to move gradually to the case where μ=1; these solutions are very different, since the average level of the capital stock is many times higher when μ=1. So, it is extremely unfair to interpret the above table as saying that one needs 2 minutes to calculate the equilibrium of the simple growth model with PEA, since this time could be greatly reduced by choosing the appropriate initial conditions. Judd [1989], for example, chooses an initial condition based on the

non-stochastic steady state, so that the initial condition already starts with a level of the capital stock that is much closer to the equilibrium one.

Finally, the speed for a given model is not very indicative of how the algorithm may perform if the model changes slightly. As we pointed out earlier, some models based on discretization and quadrature integration will suffer more from increasing the dimension of the state and the exogenous variables in the model than will PEA. Since the simple growth model has very few variables, it does not exploit all the advantages of PEA. Rather than engaging in speed races with ambiguous rules of the game, it seems more productive for researchers to generate interesting applications of the different algorithms to develop an understanding of what algorithms work better under what conditions.

A Model of Asset Prices with Endogenous Production

The representative agent maximizes the expected discounted utility as in Example 1.2 andhe can trade stocks of ownership of the (only) firm in this economy and bonds at maturities k=1,...,K ; so his budget constraint is given by

$$(3.4) \qquad c_t + s_t\, ps_t + \sum_{k=1}^{K} b_t^k\, pb_t^k = s_{t-1}\,(\, p_t + d_t\,) + \sum_{k=1}^{K} b_{t-k}^k + w_t$$

where s_t , ps_t , b_t^k , pb_t^k represent, respectively, holdings and prices of one share of ownership of the only firm in this economy, and holdings and prices of a k-period riskless real bond; d_t are the dividends distributed by the firm, and w_t is the wage received by the agent when he inelastically offers one unit of labor every period. The consumer behaves competitively in the securities, consumption and labor markets. He has one unit of labir which he supplies inelastically.

The representative firm owns the capital stock, maximizes expected discounted profits, subject to the constraints

$$c_t + i_t = f(k_{t-1}, l_t, \theta_t)$$

$$i_t = (\, k_t - \mu\, k_{t-1}\,)$$

taking prices as given and distributing the profits in the form of dividends. The productivity shock θ_t is as in Example 1.2.

If we require that

$$f(k_{t-1}, 1, \theta_t) = \theta_t \, k_{t-1}^{\alpha} \quad , \tag{3.5}$$

this means that the consumption and capital series are the same as in the model of the previous sub-section, so that the real variables are given by that model. Other than (3.5), we give ourselves enough freedom on the functional form of the production function f in order to consider different labor income and dividend processes. Clearly, w_t will be the marginal productivity of labor and $d_t = c_t - w_t$.

We know that the stock and bond prices satisfy

$$c_t^{-\gamma} \, ps_t = \delta \, E_t \left[c_{t+1}^{-\gamma} \, (ps_{t+1} + d_{t+1}) \right]$$

$$c_t^{-\gamma} \, pb_t^J = E_t \left[\delta^J \, c_{t+J}^{-\gamma} \right] \quad . \tag{3.6}$$

Once we have a solution for the real variables, these conditional expectations can be approximated by running <u>one</u> non-linear regression. Now we describe the behavior of the model with Impulse-Response functions.

<u>Impulse-Response Function of the Variables</u>

First of all we describe the behavior of the equilibrium path of consumption, investment and capital. This is important for understanding the behavior of asset prices.

Since the endogenous variables in each period depend on the state variables and these are endogenous and correlated with each other it does not make much sense to ask how each variable reacts to an increase in the state variables. We describe the behavior of the model by looking at how each variable reacts to an increase in the i.i.d. shocks to the economy, namely, the innovation to the productivity shock. At the fixed point:

$$z_t = h(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \quad ,$$

where $h: R^{\infty} \rightarrow R^n$. Although we could back out the function h implied by the approximated solution, the result would be very non-linear and hard to analyze. Instead, we consider the projection of z_t on current and past ε's

$$z_t = \sum_{l=0}^{\infty} b_l \, \varepsilon_{t-l} + \eta_t \quad ,$$

where the b's are chosen to minimize the mean square error $E\ \eta_t^2$. We can interpret η_t as the non-linear part of h . Since ε's are i.i.d. we know that

$$b_i = E [\ \varepsilon_{t-i}\ z_t\ ] / E [\ \varepsilon_t^2\ ] \quad ,$$

so that they can be easily approximated with sample covariances of the simulated series.

The coefficients $\{\ b_l^j\ \}$ scaled by $\sigma_\varepsilon / \sigma_{z_j}$ give the response of the j-th element of z_{t+i}, in terms of its standard deviation, to an increase of one standard deviation in ε_t. Besides describing how each variable responds to a shock in productivity, these coefficients will help us in explaining the covariance patterns that the model displays, because

$$(3.7) \qquad cov(\ z_{t+l}^j\ ,\ z_t^{j'}\ ) = \sigma_\varepsilon^2 \sum_{k=0}^{\infty} b_{l+k}^j\ b_k^{j'} + \eta_{t+l}^j\ \eta_t^{j'}$$

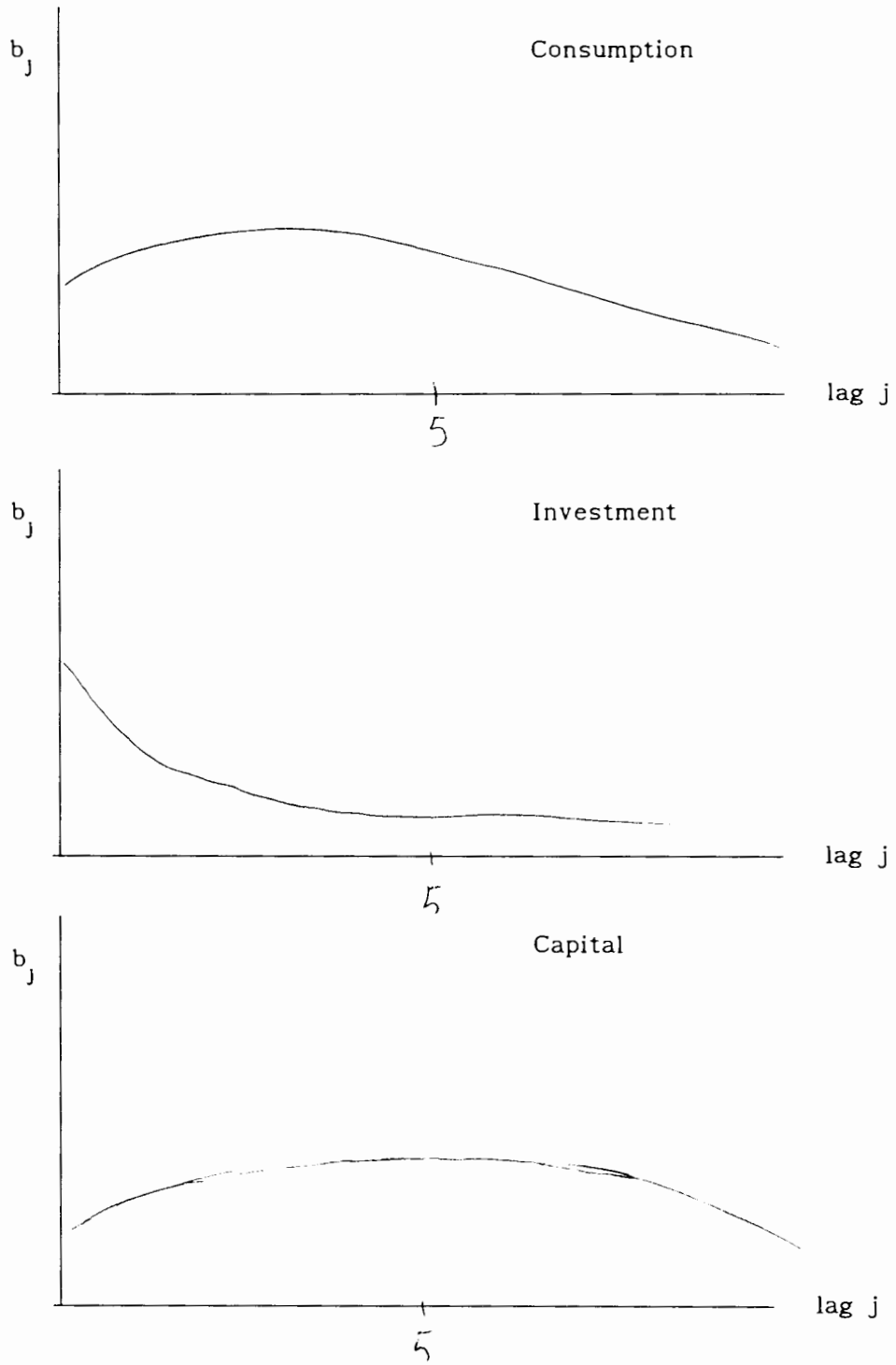
so that the convolution of the b's tells us what is the covariance, up to the non-linear terms η which, in practice, are very small.

We simulated the simple growth model with the following parameters: γ=4. , δ=.99 , μ=.98 , ρ=.95 , σ_ε=.009 and α=.33 . The choice of δ and μ suggests a quarterly model. All the calculations reported were done with 20.000 observations in the long run simulation; none of the substantive results changed when we did the same exercise with 5000 observations. We used a second order exponentiated polynomial to parameterize expectations; this did not make any difference for the real allocations but in the Euler equation for the stock price there was evidence of inaccuracy for a first degree polynomial.

The impulse-response functions for consumption, capital and GNP($\equiv c_t + k_t - \mu k_{t-1}$) are shown in figure 1. We see that the response of capital goes up for a long time, it is very permanent. The response of consumption is also hump-shaped, but it starts to decrease after four or five periods. Finally, GNP shows an exponential decrease. Let us provide some intuition for these shapes.

Figure 1

Impulse-Response for Real Variables



Clearly, investment responds much more quickly to a positive shock than consumption, so that investment absorbs most of the variation. So, after a positive shock, investment goes up immediately; this causes capital and production to be higher in the next few periods and capital is higher in the next few periods. Therefore, the hump-shaped IR for consumption is an effect of the desire for consumption smoothing that risk averse agents have and that, in a model with endogenous production, causes the volatility to be transferred to investment and not consumption. By constrast, a first order Markov process for consumption of the type that is often used in asset pricing models with exogenous production, would generate a decreasing IR function of consumption.

Bond Price Behavior and the Term Structure

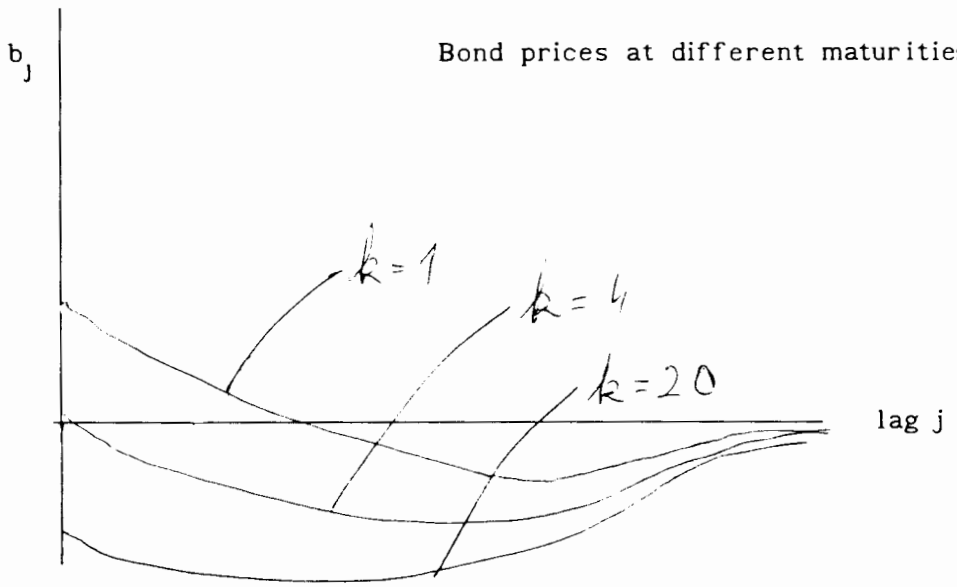
If the bond prices are given by (3.6), then the k-period interest rate is given by $1/pb_t^k$.

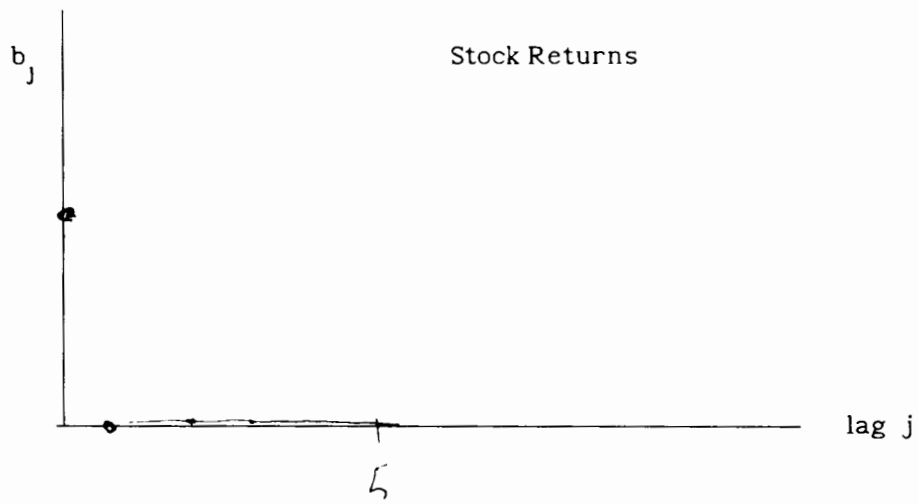
The IR of these bond prices are given in figure 2. These IR functions could have been guessed from the IR for the real variables. Consider the one-period interest rate, so that k=1; this interest rate depends only on c_t and the expectation of a function involving c_{t+1}. The IR of consumption tells us that an increase in ε_t has a bigger effect on c_{t+1} than on c_t, therefore the right side of (3.5) decreases more than the term $c_t^{-\gamma}$, so $1/pb_t^1$ responds positively to a positive shock in ε_t. However, when we consider the effect of ε_{t-j} , if j is large enough to be on the decreasing region of the consumption IR, the impact of a positive shock in lagged ε on $c_t^{-\gamma}$ is higher than the impact on $c_{t+1}^{-\gamma}$, and the interest rate responds negatively to such a shock.

For a longer-period interest rate, if k is large enough to be on the region where the IR function is decreasing, the effect of a positive shock in ε_{t-j} on c_t for any j (including j=0) is higher than the effect on c_{t+k} , so that the response of the long-term interest rate to a positive shock is negative.

Figure 2

Impulse-Response for Security Prices





This means that the response of short term interest rates to a positive innovation in productivity is opposite from the response of the long term interest. Therefore, in this model one can observe a hump-shaped term structure: if a few positive shocks in ε_t are followed by a few negative shocks, the short-term interest rates go down, while the positive and negative shocks cancel out in the long term interest rate. In this model, a humped shaped interest rate indicates a sudden drop in productivity, from better than average to worse than average. It is this sense that the Fama [1984] observation is replicated by the model.

Behavior of Stock Prices

The price of the stock ps_t depends on the specification of the production function f, which will imply a different process for d_t.

One possibility is to consider $d_t=c_t$; formally, this corresponds to the case where labor is not productive. The returns of stocks and the correlations are the same as if d_t is a fixed proportion of returns or, equivalently, if labor gets every period a fixed proportion of consumption. In this case, the price of the stock is very highly correlated with consumption, and its IR is essentially the same as that of consumption.

Under this assumption, the standard deviation of stock returns turns out to be almost the same as the standard deviation of consumption growth, which is a counterfactual implication of the model, and also one that makes it harder for the model to explain the equity premium puzzle. The equity premium in the model is of the order of .15 per cent[20], far from the 5 per cent in US data.

Another possibility for determining dividends is assuming that the production function has constant returns to scale

---

[20] We say that the risk premium is 'of the order of' because it exact value may change depending how the model is exactly simulated; these risk premia are so small that their exact value is affected by the numerical error; so using a different realization of 20.000 periods to calculate the fixed point one can find risk premia that go from .0016 to .0013 ; these differences may seem large relative to the simulated risk premium, but they do not change our conclusion that the model is far from explaining the empirical risk premium.

(3.8) $$f(k_{t-1}, l_t, \theta_t) = \theta_t \, k_{t-1}^{\alpha} \, l_t^{1-\alpha} \quad ,$$

so that the wage equals a $(1-\alpha)$ portion of the total GNP. Under this assumption we know that $k_t=p_t$.

Relative to the previous scheme where $c_t=d_t$, now the wage absorbs a larger part of the randomness of the economy, and the implied dividend process is less volatile than before. Then the standard deviation of stock returns is about .0019, almost one-tenth the standard deviation of GNP growth. Therefore, this model goes in the opposite direction: it makes stock returns much less volatile than GNP growth. The risk premium is almost exactly zero; in fact it turns out that not only the mean, but also the standard deviation of bond and stock returns are equal.

Finally, we consider the case where the labor income is constant. In this way we cause all the variance in consumption to be totally absorbed by dividends; as we increase the fixed wage the variance of the dividend stays constant but its mean goes down, so that the standard deviation as a proportion of the mean can become arbitrarily high. This can be justified by the following production function

(3.9) $$f(k_{t-1}, l_t, \theta_t) = \theta_t \, k_{t-1}^{\alpha} + \zeta \, (l_t - 1) \quad ,$$

and a perfectly competitive labor market as well as by various models with imperfect labor markets. The production function (3.9) has constant returns to scale on labor and decreasing returns to scale in capital, so there are decreasing returns to scale overall and the maximization problem of the firm is well defined. In this model, $w_t=\zeta$ and $d_t= c_t-\zeta$.

We choose a level of wage that delivers standard deviation in stock returns of .04, . The correlation coefficient of stock returns and ε is .97 so that, as with consumption growth, the impulse response function is concentrated at zero. Even with so much volatility, however, the risk premium is only .32 per cent. This means that even after increasing the volatility of stock returns we can not explain the risk premium in the data.

Time-Dependence of Volatility in Stock Returns

As pointed out by Rowenhorst, the model matches the empirical evidence suggesting that volatility of stock returns is higher in recessions. We would like to argue that, in fact, this is true of many equilibrium models, and it is not particularly a result that depends

on the endogeneity of production.

In most equilibrium asset pricing models, including the one in this paper, the stock price is almost perfectly correlated with dividends, and the stock return is very highly correlated with the innovation to production. This can be checked by simulation and by several analytic solutions that are available. In this case, $p_t \cong K\ d_t$ for some constant K. Also, assume that the innovation of dividends has a constant variance σ_η^2. Then, the conditional variance of the stock return is given by

$$E_t\left[\frac{p_{t+1}+d_{t+1}}{d_t} - E_t\left[\frac{p_{t+1}+d_{t+1}}{d_t}\right]\right]^2 \cong (1/d_t)^2\ (K+1)^2\ \sigma_\eta^2 \quad .$$

Clearly this is large when dividends are small, and in equilibrium models dividends are small in recessions. Since the stock return is practically i.i.d. (by virtue of being almost perfectly correlated with the innovations to the exogenous processes), this implies that volatility will be higher in recessions.

Conclusion

In this paper we have introduced a method for solving nonlinear, stochastic dynamic models with rational expectations and we have applied it to the analysis of asset prices in a production economy. The solution procedure can get around many problems that appear often in simulation: it works from the Euler equations, equilibrium conditions and budget constraints, so that there is no need to have a planner's problem describing the equilibrium and it can approximate non-linearities arbitrarily well. The model is flexible enough to easily accomodate models with distortions, heterogeneous agents, inequality constraints and discrete choice.

Many of the computational barriers that are faced by other methods are not immediately binding here: the computational costs do not grow exponentially with the number of state variables; because instead of using quadrature as a method of integration we use Monte-Carlo integration we can easily handle a large number of shocks in the economy. There is already a large number of applications of this algorithm to different types of models.

In this paper we have concentrated on discussing the applicability of the method and its validity, and we have not discussed ways of increasing the computational speed; research on this area is being done at present. The algorithm can be improved in several directions: the fixed point β_f could be calculated with a number of techniques, the expectations involved in $S(\beta)$ can be approximated more efficiently using techniques for fast Monte-Carlo integration, there should be classes of functions other than the polynomials used in this paper that can capture better the non-linearities in particular models. Speed can be improved byrunning the non-linear regressions in only one step (in the fixed point this will not matter), by using in the regression only a sampled process and by efficient ways of calculating the derivative of S. However, one of the advantadges of the parameterized expectations approach is its great simplicity, and it is likely that these improvements will complicate considerably the application of the algorithm.

One of the main advantages of PEA is that it endogenously selects the region where the policy function is approximated and that, at the same time, it performs Monte-Carlo integration, so that relatively large models can be easily handled.

Other extensions of the algorithm are designed to handle problems where there is no steady-state distribution or the researcher is interested in looking at the model away from the steady state. In this case, the long run simulations of Step 3 should be replaced by repeated short run simulations, as in Marshall [1988] and Marcet and Marimon [1991]. Another extension that has not yet been pursued is to models with private information; it is possible to imagine how the algorithm could be modified to handle this case, where additional

lags, or functions of the observed variables would be introduced according to their predictive power.

In our analysis of asset pricing in a production economy we saw that the model could generate an interesting pattern of stock and bond prices. With a very simple growth model it was easy to generate asset prices with second order properties very different from the ones in the usual model when consumption is exogenous, equal to dividends and given by a Markov process of order one, and it was possible to generate a humped-shaped term structure of interest rates in recessions. The reason for the hump-shape was the possibilities and the desire for consumption smoothing present in a model with investment and risk averse agents. This is important because some researchers may have been driven away form this model thinking that it imposed too many restrictions on the data, so that its empirical performance was bound to be even worse than that of exogenous consumption. This is one of those conforting developments in economics where better performance is obtained by imposing more structure in the model instead of the usual approach of making the model more general and increasing the number of parameters.

We also argued that the evolution of the volatility of the stock returns across time matched the one observed in the data, and that the risk premium was higher than usual, but still far away from the observed in postwar US data. Contrary to previous papers, we showed how the higher volatility of stock returns in recessions is easily reproduced in most equilibrium asset pricing models, so both of these results do not differ much from others that can be obtained with exogenous production.

References

Backus, D.K., Gregory, A.W. and Zin S. (1989); "Risk Premia in the Term Structure: Evidence from Artificial Economies" Journal of Monetary Economics, 24

Bansal, R. (1988); "Boundedly Rational Learning", manuscript, Carnegie-Mellon University.

Baxter, M. (1988); "Approximation Suboptimal Dynamic Equilibrium: An Euler Equation Approach" University of Rochester working paper.

Bizer, D. and K. Judd (1989); "Taxation under Uncertainty" American Economic Review, May.

Brock, W.A. (1982) "Asset Prices in a Production Economy", in J.J McCall, ed. The Economics of Information and Uncertainty, University of Chicago Press.

Brock, W. A., L. J. Mirman (1972); "Optimal Economic Growth and Uncertainty: the Discounted Case", J.E.T., vol. 4, 479-513.

Coleman, W.J. (1989) "Equilibrium in a Production Economy with an Income Tax", working paper, Board of Governors of the Federal Reserve System.

den Haan, W. (1990); "Optimal Inflation in a Sidrausky Model" Journal of Monetary Economics.

den Haan, W. (1990); "The Term Structure of Interest Rates in Real and Monetary Production Economies" working paper, Carnegie Mellon University.

den Haan, W. and Marcet A. (1989); "Accuracy in Simulations" working paper Carnegie Mellon University.

den Haan and Marcet (1990); "Solving a Simple Growth model by Parameterizing Expectations" JBES, forthcoming.

Epstein, L.G. and Zin S.E. (1989); "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns II", working paper, Carnegie Mellon University.

Fair, R. C., and J. B. Taylor (1983); "Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models", Econometrica, vol. 51, 1169-1185.

Fama, E.F. (1984) "The Information in the Term Structure" Journal of Financial Economics, 25.

Garcia, C. B. and W. I. Zangwill; Pathways to Solutions, Fixed Points, and Equilibria. Prentice Hall.

Hansen, L. P. (1982); "Large Sample Properties of Generalized Method of Moments Estimators", Econometrica, vol. 50, 1029-1054.

Heaton, J. (1990); "The Interaction Between Time Non-Separable Preferences and Time Aggregation", working paper, Massachussetts Institute of Technology.

Ingram, Beth (1986); "Equilibrium Modelling of Asset Prices: Rationality Versus Rules of Thumb", manuscript, Cornell University.

Judd, K. (1989); "Minimum Weighted Residual Methods in Dynamic Economic Models" working paper, Hoover Institution.

Ketterer and Marcet (1989); "Introducing Derivative Securities: A General Equilibrium Approach". Preliminary Draft, Carnegie Mellon U.

Kydland, F. E. and E. C. Prescott (1982); "Time to Build and Aggregate Fluctuations", Econometrica, vol. 50, 1345-1370.

Labadie, P. (1986); "Solving Non-Linear Rational Expectations Models for Estimation: An Application to an Asset Pricing Model with Currency", manuscript, Columbia University.

Ljung, L. and T. Söderström (1983); Theory and Practice of Recursive Identification, The MIT Press.

Lucas, R. E. (1978); "Asset Prices in an Exchange Economy", Econometrica, vol. 46, 1426-1446.

Marcet, A. (1988); "Solving Non-linear Stochastic Models by Parameterizing Expectations", working paper, Carnegie Mellon University.

Marcet, A. and Marimon, R. (1991); "Communication, Commitment and Growth", working paper, Universitat Pompeu Fabra.

Marcet, A. and T. J. Sargent (1989a); "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", Journal of Economic Theory, August.

Marcet, A. and T. J. Sargent (1989b); "Convergence of Least Squares Learning in Environments with Hidden State Variables and Private Information", Journal of Political Economy , December.

Marcet and Singleton (1990) "Equilibrium Asset Prices and Savings of Heterogeneous Agents with Protfolio Constraints" Working Paper.

Marcet, A. (1991) "Simulation Analysis of Stochastic Dynamic Models: Applications to Theory and Estimation", Invited Symposium of the VI World Congress of the Econometric Society. Forthcoming in Advances in

Econometrics, ed. by J.J Laffont and C.A. Sims.

Marshall, D. (1988); "Inflation and Asset Returns in a Monetary Economy with Transaction Costs". Working Paper, Northwestern University.

Mehra, R. and E. C. Prescott (1985); "The Equity Premium: A Puzzle", Journal of Monetary Economics.

Miller R. (1984); "Job Matching and Occupational Choice", Journal of Political Economy , vol. 92, 1086-1120.

Novales, A. (1990):"A Stochastic Monetary Equilibrium Model of the Interest Rate" Econometrica

Pakes, A. (1986); "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks", Econometrica, vol. 54, 755-784.

Riesz, T.A. (1988); "The Equity Risk Premium: A Solution" Journal of Monetary Economics 22

Rojas, G. (1991); "Taxation in a Model with Public Capital", working paper, Universitat Pompeu Fabra.

Rouwenhorst, K.G. (1989) "Asset Returns and Business Cycles: A General Equilibrium Approach". University of Rochester working paper.

Rust, J. (1987); "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher", Econometrica, vol. 55, 999-1033.

Schwert, G.W. (1989); "Business Cycles, why does Stock Market Volatility Change Over Time?" Journal of Finance, 44

Sims, C. A. (1985); "Solving Non-Linear Stochastic Equilibrium Models 'Backwards'", manuscript, University of Minnesota.

Tauchen, G. (1986); "Quadrature-Based Methods for Obtaining Approximate Solutions to the Integral Equations of Nonlinear Rational Expectations Models", manuscript, Duke University.

Taylor, J. and H. Uhlig, (1990); "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods" Journal of Business Economics and Statistics, January.

Townsend, R. (1983); "Forecasting the Forecasts of Others", Journal of Political Economy.

Wolpin, K. I. (1984); "An Estimable Dynamic Stochastic model of Fertility and Child Mortality", Journal of Political Economy, vol. 92, 852-874.

RECENT WORKING PAPERS

1. ALBERT MARCET and RAMON MARIMON, *Communication, Commitment and Growth* (June 1991).

2. ANTONI BOSCH, *Economies of Scale, Location, Age, and Sex Discrimination in Household Demand* (June 1991).

3. ALBERT SATORRA, *Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures* (June 1991).

4. JAVIER ANDRES and JAUME GARCIA, *Wage Determination in the Spanish Industry* (June 1991).

5. ALBERT MARCET, *Solving Non-Linear Stochastic Models by Parameterizing Expectations: An Application to Asset Pricing with Production* (July 1991).

UNIVERSITAT POMPEU FABRA
Balmes, 132
Telephone (93) 484 97 00
Fax (93) 484 97 02
Barcelona 08008