Algorithms for Solving Dynamic Models with Occasionally Binding Constraints*

Lawrence J. Christiano† Jonas D.M. Fisher‡

September, 1998

Abstract

We describe and compare several algorithms for approximating the solution to a model in which inequality constraints occasionally bind. Their performance is evaluated and compared using various parameterizations of the one sector growth model with irreversible investment. We develop parameterized expectation algorithms which, on the basis of speed, accuracy and convenience of implementation, appear to dominate the other algorithms.

JEL Classification: C6, C63, C68
Keywords: Occasionally binding constraints, Parameterized Expectations, Collocation, Chebyshev, Interpolation.

*We thank Graham Candler, Wouter den Haan, Kenneth Judd, Albert Marcet, David Marshall, Mario Miranda, Ellen McGrattan, Manuel Santos, two anonymous referees and an associate editor for helpful comments. Christiano is grateful for the financial support of a grant to the National Bureau of Economic Research from the National Science Foundation. Fisher is grateful for the financial support of the Social Sciences and Humanities Research Council. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

†Northwestern University, NBER, Federal Reserve Bank of Chicago.

‡(Corresponding author) Research Department, Federal Reserve Bank of Chicago, 230 South LaSalle Street, Chicago, IL 60604-1413. Voice: (312) 322-8177, Fax: (312) 322-2357.

1. Introduction

There is considerable interest in studying the quantitative properties of dynamic general equilibrium models. For the most part, exact solutions to these models are unobtainable and so in practice researchers must work with approximations. An increasing number of the models being studied have inequality constraints that occasionally bind. Important examples of this are heterogeneous agent models in which there are various kinds of constraints on the financial assets available to agents.[1] Other examples include multisector models with limitations on the intersectoral mobility of factors of production, and models of inventory investment.[2] For researchers selecting an algorithm to approximate the solution to models like these, important criteria include numerical accuracy and programmer and computer time requirements. That the last of these should remain a significant consideration is perhaps surprising, in view of the dramatic, ongoing pace of improvements in computer technology. Still, the economic models being analyzed are growing in size and complexity at an even faster pace, and this means that efficiency in the use of computer time remains an important concern in the selection of a solution algorithm. Our purpose in this paper is to provide information useful to researchers in making this selection.

We describe several algorithms, and evaluate their performance in solving the one-sector infinite horizon optimal growth model with a random disturbance to productivity. In this model the nonnegativity constraint on gross investment is occasionally binding. We chose this model for two reasons. First, its simplicity makes it feasible for us to solve the model by doing dynamic programming on a very fine capital grid. Because we take the dynamic programming solution to be virtually exact, this constitutes an important benchmark for evaluating the algorithms considered. Second, the one sector growth model is of independent interest, since it is a basic building block of the type of general equilibrium models analyzed in the literature.[3]

All the methods we consider work directly or indirectly with the Euler equation associated with the recursive representation of the model, in which the nonnegativity constraint is accommodated by the method of Lagrange multipliers. Suitably modified versions of the algorithms emphasized by Bizer and Judd (1989), Coleman (1988), Danthine and Donaldson (1981) and

Judd (1992) work directly with this formulation, and are evaluated here. We also consider the algorithm advocated by McGrattan (1996), in which the multiplier and policy function are approximated using an approach based on penalty functions. Finally, we consider an algorithm in which policy and multiplier functions are approximated indirectly by solving for the conditional expectation of next period's marginal value of capital. Algorithms which solve for a conditional expectation function in this way are referred to as parameterized expectations algorithms (PEA). We describe PEAs which are at least as accurate as all the other algorithms considered and which dominate these other algorithms in terms of programmer and computer time.[4]

The first use of a PEA appears to be due to Wright and Williams (1982a, 1982b, 1984), and was further developed by Miranda (1985), and Miranda and Helmberger (1988). Later, a variant on the idea was independently discovered in the macroeconomics literature by Marcet (1988). Marcet's PEA, which we call conventional PEA, has been applied to a great variety of interesting problems, many of which are cited in Marcet and Marshall (1994).

An important advantage of PEAs when there are occasionally binding constraints is that they make it possible to avoid a cumbersome direct search for the policy and multiplier functions that solve the Euler equation. Methods which focus on the policy function must jointly parameterize these two functions, and doing this in a way that the Kuhn-Tucker conditions are satisfied is tricky and adds to programmer time. An alternative to working with Lagrange multipliers is to work with a penalty function formulation. However, this approach requires searching for the value of a penalty function parameter, and this can add substantially to programmer and computer time. PEAs exploit the fact that Euler equations and Kuhn-Tucker conditions imply a convenient mapping from a parameterized expectation function into policy and multiplier functions, eliminating the need to separately parameterize the latter. In addition, the search for a conditional expectation function that solves the model can be carried out without worrying about imposing additional side conditions analogous to the Kuhn-Tucker conditions. In effect, by working with a PEA one reduces the number of unknowns to be found, and eliminates a set of awkward constraints.

Alternative PEAs differ on at least two dimensions: the particular conditional expectation function being approximated and the method used to compute the approximation. Marcet's

approach approximates the expectation conditional on the beginning of period state variables, while Wright and Williams propose approximating the expectation conditional on a current period decision variable. A potential shortcoming of Marcet's approach is that functions of beginning of period state variables tend to have kinks when there are occasionally binding constraints. The conditional expectation function that is the focus of Wright and Williams' analysis, by contrast, appears to be smoother in the growth model. Being smoother, the Wright and Williams conditional expectation is likely to be easier to approximate numerically. This deserves further study in other applications.

We also describe improvements on the conventional PEA's method for approximating the conditional expectation. A key component of conventional PEA is a cumbersome nonlinear regression step, potentially involving tens of thousands of synthetic data points. We show that such a large number of observations is required because the conventional PEA inefficiently concentrates the explanatory variables of the regression in a narrow range about the high probability points of the invariant distribution of a model. This feature of the method is sometimes cited as a virtue in the analysis of business cycle models, where one is interested in characteristics of the invariant distribution such as first and second moments. But, it is well known in numerical analysis that the region where one is interested in a good quality fit and the region one chooses to emphasize in constructing the approximation need not coincide.

This point plays an important role in our analysis and so it deserves emphasis. A classic illustration of it is based on the problem of approximating the function, $1/(1+k^2)$, defined over $k \in [-5,5]$, by a polynomial.[5] If one cared uniformly over the domain about the quality of fit, then it might seem natural to select an equally spaced grid in the domain and choose the parameters of the polynomial so that the two functions coincide on the grid points. But, it is well known that this strategy leads to disaster in this example. The upper panel in Figure 1 shows that the 10^{th} order polynomial approximating function exhibits noticeable oscillations in the tails when this method is applied with the 11 grid points indicated by +'s in the figure. Moreover, when more grid points are added, keeping the distance between grid points constant, the oscillations in the tail areas become increasingly violent without bound. Not surprisingly, one way to fix this problem is to redistribute grid points a little toward the tail areas. This is

what happens when the grid points are chosen based on the zeros of a Chebyshev polynomial, as in the lower panel in Figure 1. Note how much better the approximation is in this case. In addition, it is known that as the number of grid points is increased, the approximating function converges to the function to be approximated in the sup norm. Thus, even if one cares uniformly over the interval about the quality of fit, it nevertheless makes sense to 'oversample' the tail areas, using the zeros of a Chebyshev polynomial.

According to the Chebyshev interpolation theorem, this is a very general result and holds for a large class of functions, not just the one depicted in Figure 1. We take our cue from this in modifying conventional PEA so that tail areas receive relatively more weight in approximating the model's solution.[6] As a result, we are able to get superior accuracy with many fewer synthetic data points (no more than 10!). Moreover, the changes we make convert the non-linear regression of conventional PEA into a linear regression with orthogonal explanatory variables. The appendix to Christiano and Fisher (1994), which shows how to implement our procedure in a version of the growth model with an arbitrary number of capital stocks and exogenous shocks, establishes that the linearity and orthogonality property generalizes to arbitrary dimensions. In this paper, we show that when applied to the standard growth model, our approach produces results at least as accurate as the best other method considered, and is orders of magnitude faster.[7]

The paper is organized as follows. In the following section the model to be solved is described, and various ways of characterizing its solution are presented. The algorithms discussed later make use of these alternative characterizations. In section 3 we describe a general framework which contains the algorithms considered here as special cases. Having a single framework is convenient both for presenting and comparing the various solution methods considered. Section 4 presents a discussion of a subset of the algorithms considered. The details of other algorithms appear in an appendix to this paper. Section 5 presents the results of our numerical analysis of case studies. The final section offers some concluding remarks.

2. The Model and Alternative Characterizations of Its Solution

In this section we present the model that we study, and we provide three alternative characterizations of its solution. The first exploits the Lagrangian formulation of our problem, and characterizes a solution by a policy and a multiplier function. The second and third characterize the solution as two different conditional expectation functions. Although the first of these characterizations is well known, we discuss it nevertheless in order to set up notation and concepts useful for discussing the second and third characterizations. A fourth characterization, based on penalty functions, is described in the appendix. The four characterizations form the basis for the computational algorithms studied in this paper.

After discussing the model and alternative characterizations of its solution, we present the formulas that we use to compute asset prices and rates of return for our model economy.

2.1. The Model

We study a simple version of the stochastic growth model. At date 0 the representative agent values alternative consumption streams according to $E_0 \sum_{t=0}^{\infty} \beta^t U(c_t)$, where c_t denotes time t consumption, $\beta \in (0,1)$ is the agent's discount factor, and U denotes the utility function. The aggregate resource constraint is given by

$$c_t + \exp(k_{t+1}) - (1-\delta)\exp(k_t) \leq f(k_t, \theta_t) \equiv \exp(\theta_t + \alpha k_t), \tag{1}$$

where $k_t \in [\underline{k}, \overline{k}] \subset \Re$ denotes the logarithm of the beginning-of-period-t stock of capital, and δ, $\alpha \in (0,1)$. Here, $-\infty < \underline{k} < \overline{k} < \infty$, δ is the rate of depreciation of capital, and α is capital's share in production. We assume $\theta_t \in \Theta$ has a first order Markov structure with the density of θ_{t+1} conditional on θ_t given by $p_{\theta'}(\theta_{t+1} \mid \theta_t)$. In the irreversible investment version of the model, we require that gross investment be non-negative, *i.e.*:

$$\exp(k_{t+1}) - (1-\delta)\exp(k_t) \geq 0. \tag{2}$$

In the reversible investment version, (2) is ignored.

2.2. The Solution as Policy and Lagrange Multiplier Functions

Let $h(k,\theta)$ denote the Lagrange multiplier on (2) in the planning problem associated with this model economy. According to one characterization, the solution to the planning problem is a set of time invariant functions $g : [\underline{k}, \overline{k}] \times \Theta \to [\underline{k}, \overline{k}]$, and $h : [\underline{k}, \overline{k}] \times \Theta \to \Re_+$ satisfying an Euler equation,

$$R(k,\theta; g, h) = 0, \text{ for all } (k,\theta) \in [\underline{k}, \overline{k}] \times \Theta, \tag{3}$$

and a set of Kuhn-Tucker conditions

$$\exp\left(g(k,\theta)\right) - (1-\delta)\exp(k) \geq 0, \;\; h(k,\theta) \geq 0, \text{and} \;\; h(k,\theta)\left[\exp\left(g(k,\theta)\right) - (1-\delta)\exp(k)\right] = 0, \tag{4}$$

for all $(k,\theta) \in [\underline{k}, \overline{k}] \times \Theta$. Here,

$$R(k,\theta; g, h) = U_c\left(k, g(k,\theta), \theta\right) - h(k,\theta) - \beta \int m\left(g(k,\theta), \theta'; g, h\right) p_{\theta'}(\theta' \mid \theta) d\theta', \tag{5}$$

and

$$m(k', \theta'; g, h) = U_c\left(k', g(k',\theta'), \theta'\right)\left[f_k\left(k', \theta'\right) + 1 - \delta\right] - h\left(k', \theta'\right)(1-\delta) \geq 0. \tag{6}$$

In (6), $f_k = \alpha \exp\left(\theta' + (\alpha - 1)k'\right)$ is next period's marginal product of capital, given that the beginning-of-next-period's capital stock is k' and next period's technology shock is θ'. Similarly, $U_c(k', g(k',\theta'), \theta')$ is next period's marginal utility of consumption, given that next period's investment decision is determined by g, and that consumption has been substituted out using (1) evaluated at equality. In (5), the beginning-of-next-period's capital stock, k', is evaluated under the assumption that this period's investment decision is determined by g.

The inequality in (6) reflects: (i) m is the derivative of the value function associated with the dynamic programming formulation of the planning problem, and (ii) a suitably modified version of the proof to Theorem 4.7 in Stokey and Lucas, with Prescott (1989) can be used to show that the value function is increasing in the capital stock.[8] Thus, one way to characterize a solution to the model is that it is a pair of functions, g and h, that are consistent with the Kuhn-Tucker conditions and that also satisfy the functional equation, $R(k,\theta; g, h) = 0$.[9]

2.3. The Solution as a Conditional Expectation Function

Solutions to the growth model can also be characterized in terms of various conditional expectation functions. We first discuss the conditional expectation that is the focus of Marcet (1988)'s analysis and we then consider the conditional expectation used by Wright and Williams (1982a, 1982b, 1984).

2.3.1. A Characterization Due to Marcet

According to the approach used by Marcet (1988), a solution is a function, $e : [\underline{k}, \overline{k}] \times \Theta \longrightarrow \Re$ satisfying

$$\bar{R}^{pea}(k, \theta; e) = 0, \text{ for all } (k, \theta) \in [\underline{k}, \overline{k}] \times \Theta, \tag{7}$$

where

$$\bar{R}^{pea}(k, \theta; e) = \exp[e(k, \theta)] - \int m\left(g(k, \theta), \theta'; g, h\right) p_{\theta'}(\theta' \mid \theta) d\theta', \tag{8}$$

and m is defined in (6). Evidently, $\exp[e(k, \theta)]$ is a conditional expectation function. The functions g and h on the right of the equality in (8), are derived from e. To see how, first let the function $\bar{g} : [\underline{k}, \overline{k}] \times \Theta \longrightarrow \Re$ be defined implicitly by:

$$U_c\left(k, \bar{g}(k, \theta), \theta\right) = \beta \exp[e(k, \theta)].$$

Then,

$$g(k, \theta) = \begin{cases} \bar{g}(k, \theta), \text{ if } \bar{g}(k, \theta) > \log(1 - \delta) + k \\ \log(1 - \delta) + k, \text{ otherwise,} \end{cases} \tag{9}$$

and

$$h(k, \theta) = U_c\left(k, g(k, \theta), \theta\right) - \beta \exp[e(k, \theta)]. \tag{10}$$

This mapping guarantees that g and h satisfy the Kuhn-Tucker conditions, regardless of the choice of function, $e : [\underline{k}, \overline{k}] \times \Theta \longrightarrow \Re$. To see this, note first that, trivially, $\bar{g}(k, \theta) \geq \log(1 - \delta) + k$ implies $h(k, \theta) = 0$. Also, if $\bar{g}(k, \theta) < \log(1 - \delta) + k$, then $h(k, \theta) > 0$ because of the strict concavity of the utility function.

For computational purposes, it is useful to note that the e function which solves the model can equivalently be characterized as satisfying:

$$R^{pea}(k,\theta;e) = 0, \text{ for all } (k,\theta) \in [\underline{k},\overline{k}] \times \Theta, \tag{11}$$

where

$$R^{pea}(k,\theta;e) = e(k,\theta) - \log\left[\int m\left(g(k,\theta),\theta';g,h\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right], \tag{12}$$

and m, g, and h are defined according to (6), (9) and (10), respectively.

2.3.2. A Characterization Due to Wright and Williams

Wright and Williams (1982a, 1982b, 1984) work with a slightly different conditional expectation function. Instead of expressing it as a function of the current period state variables, k, θ, they express it as a function of the current period decision, k', and the current period technology shock, θ. In particular, they characterize a solution as a function $v : [\underline{k},\overline{k}] \times \Theta \longrightarrow \Re$, satisfying

$$\tilde{R}^{pea}(k',\theta;v) = 0, \text{ for all } (k',\theta) \in [\underline{k},\overline{k}] \times \Theta, \tag{13}$$

where

$$\tilde{R}^{pea}(k',\theta;v) = v(k',\theta) - \log\left[\int m\left(k',\theta';g,h\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right], \tag{14}$$

and m is defined in (6). To construct m, we need next period's policy and multiplier functions, g and h. These are derived from v as follows. First, let the function $\bar{g} : [\underline{k},\overline{k}] \times \Theta \longrightarrow \Re$ be defined implicitly by:

$$U_c\left(k',\bar{g}(k',\theta'),\theta'\right) = \beta \exp\left[v\left(\bar{g}(k',\theta'),\theta'\right)\right].$$

Then, g is defined by (9) with k, θ replaced by k', θ' and h is defined by

$$h(k',\theta') = U_c\left(k',g(k',\theta'),\theta'\right) - \beta \exp\left[v(g(k',\theta'),\theta')\right]. \tag{15}$$

With the above operator from v to g and h, the Kuhn-Tucker conditions are *not* satisfied for

arbitrary $v : [\underline{k}, \overline{k}] \times \Theta \longrightarrow \Re$. In particular, for a v function that is sufficiently increasing in its first argument, $\bar{g}(k, \theta) < \log(1-\delta)+k$ implies $h(k, \theta) < 0$. Moreover, a sufficiently non-monotone v function could imply a g that is a correspondence rather than a function. These may not be problems in practice. First, it is easily verified that for v functions which are decreasing in their first argument, the above operator does guarantee that the Kuhn-Tucker conditions are satisfied and that g is a function. Second, concavity of the value function and the fact that m is the derivative of the value function with respect to capital, implies that the exact v function is decreasing in its first argument. Third, an operator useful in computing v, which maps the space of functions $v : [\underline{k}, \overline{k}] \times \Theta \longrightarrow \Re$ into itself, has the property of mapping the subspace of v functions decreasing in k' into itself. For an arbitrary v this operator, $\mathcal{P}(v)$, is defined as follows:

$$\mathcal{P}(v)(k', \theta) = \log \left[\int m\left(k', \theta'; g, h\right) p_{\theta'}(\theta' \mid \theta) d\theta' \right], \tag{16}$$

where g and h are obtained from v in the way described above.[10] Thus, as long as it begins with a v function decreasing in k', an algorithm that approximates v as the limit of a sequence of functions generated by the $\mathcal{P}$ operator may never encounter v functions which imply g and h that are not functions or are inconsistent with the Kuhn-Tucker conditions. Still, with other model economies and other types of computational algorithms one clearly has to be on the alert for these possibilities. We investigate them in the numerical analysis below.

2.4. Discussion

It is easily confirmed that the solutions to the four functional equations, R, R^{pea}, $\bar{R}^{pea}$, and $\tilde{R}^{pea}$, correspond to four equivalent characterizations of the solution to the model. From a computational perspective, however, they are quite different when (2) binds occasionally. A computational strategy based on solving the functional equation, $R = 0$, requires finding two functions, g and h, subject to the constraint that they satisfy the Kuhn-Tucker conditions. In contrast, finding e to solve $R^{pea} = 0$, $\bar{R}^{pea} = 0$, or v to solve $\tilde{R}^{pea} = 0$ involves having to find only one function. Moreover, strategies based on finding e need not impose any extra side conditions. Finally, an argument presented above as well as numerical results reported below suggest that

in practice this may be true for v as well.

There are some additional differences between the characterizations based on e and v. First, in our model economy the operator from e to g and h has a closed form expression and so is trivial to implement computationally. In contrast, the analogous operator from v to g and h requires solving a nonlinear equation, and so is computationally more burdensome. This distinction *per se* is not particularly significant, however, since it reflects a special feature of our model economy. In general the mapping from e to g and h also requires solving a nonlinear equation.

A potentially more important difference is that v is likely to be smoother than e. This makes v an easier object to approximate numerically. By v being smoother than e we mean that v is more likely to be differentiable in its two arguments than is e. To understand this, it is useful to recall that these functions must satisfy

$$\begin{aligned} v(k',\theta) &= \log\left[\int m\left(k',\theta';g,h\right)p_{\theta'}(\theta' \mid \theta)d\theta'\right], \\ e(k,\theta) &= \log\left[\int m\left(g(k,\theta),\theta';g,h\right)p_{\theta'}(\theta' \mid \theta)d\theta'\right] \equiv v\left(g(k,\theta),\theta\right), \end{aligned} \tag{17}$$

where m is defined in (6). These relations indicate that the smoothness properties of e and v depend on the smoothness properties of g and h. We first consider smoothness in k and we then consider smoothness in θ.

When viewed as functions of k, for fixed θ, g and h are expected to have a kink (i.e., non-differentiability) at the value of the capital stock where the irreversibility constraint begins to bind. We expect this kink to have a greater impact on e than on v. To see why, fix θ and consider the case where Θ contains a finite number of l elements, $\{\theta_1,...,\theta_l\}$, each having positive probability. Then, (17) reduces to

$$e(k,\theta) = \log\left[\sum_{i=1}^{l} m\left(g(k,\theta),\theta_i;g,h\right)p_{\theta'}(\theta_i \mid \theta)\right].$$

From this expression, we can see that there are two types of kink in e. The first type occurs at the value of k where the irreversibility constraint binds in the current period. This operates

through the kink in g, and has a relatively large impact on e because it is present in each term in the summand. The second type of kink occurs for values of k which cause the irreversibility constraint to bind in the next period for some $\theta' \in \Theta$. This type of kink is likely to have a relatively small impact on e because it affects only one of the terms in the summand.[11] Similarly, when the distribution of θ' has continuous support, then, as long as the distribution does not have mass points, we expect the second type of kink in e to vanish altogether. To summarize, whether the distribution of θ' is continuous or discrete, we expect the impact on e of the first type of kink to be relatively large and of the second type of kink to be small or non-existent.

By contrast, the function v does not have the first type of kink because the current period policy rule simply does not enter v. We expect v to be a relatively smooth function in k' for fixed θ because this function has only the second type of kink. In the case where θ' has a continuous distribution, we expect v to have no kinks in k' at all.[12]

We now consider the smoothness of v and e in θ for each fixed k. We consider this issue in the only case where it is interesting, namely, where θ has continuous support. First, the only way that θ enters v is via the conditional density, $p_{\theta'}(\theta' \mid \theta)$. The conditional densities typically used in practice are differentiable in θ, and so for these densities, v is too.[13] Second, when g is viewed as a function of θ for fixed k, it is likely to have a kink in it. The argument in the previous paragraphs indicates that this kink will be inherited by the e function too. We conclude that v is also likely to be smoother than e in θ.

2.5. Asset Prices

We are interested in the properties of the quantity allocations that solve the planning problem, and also in the rates of return and prices in the underlying competitive decentralization. In particular, we are interested in the consumption cost of end-of-period capital (i.e., Tobin's q) and the annualized rate of return on equity and risk free debt, R^e and R^f. We think of the time unit of the model as being one quarter, so that obtaining annualized returns requires raising quarterly returns to a fourth power. The rates of return and prices that interest us are:

$$q(k,\theta) \quad = \quad 1 - \frac{h(k,\theta)}{U_c\left(k, g(k,\theta), \theta\right)}, \tag{18}$$

$$R^f(k,\theta) = 100\left\{\left[\frac{U_c\left(k,g(k,\theta),\theta\right)}{\beta\int U_c\left(g(k,\theta),g\left(g(k,\theta),\theta'\right),\theta'\right)p_{\theta'}(\theta'\mid\theta)d\theta'}\right]^4-1\right\},$$

$$R^e(k,\theta,\theta') = 100\left\{\left[\frac{f'\left(g(k,\theta),\theta'\right)+(1-\delta)\,q\left(g(k,\theta),\theta'\right)}{q(k,\theta)}\right]^4-1\right\}.$$

It is easy to establish that $0 \leq q(k,\theta) \leq 1$. The result, $q \leq 1$, follows from the non-negativity of the Lagrange multiplier, h. The result, $q \geq 0$, follows from (3), (5), $U_c \geq 0$, and the non-negativity of m in (6). The event in which the constraint binds corresponds to the event $q(k,\theta) < 1$. It is easily verified that in a competitive decentralization of this economy where households own the capital stock and undertake investment, q is the price of end-of-period capital in consumption units, R^e is the rate of return on capital, and R^f is the rate of return on risk free debt.[14]

3. Weighted Residual Solution Methods

The computational algorithms we consider in this paper are special cases of the framework in Reddy (1993)'s numerical analysis text, which corresponds closely to the framework presented in Judd (1992, 1998). This framework is designed for problems in which one seeks a function, say $f : D \to Q$, which solves the functional equation, $F(s;f) = 0$ for all $s \in D$, where D is a compact set. This can be a difficult problem when, as in our case, there is a continuum of elements in D. Then, finding a solution corresponds to a problem of solving a continuum of equations (one for each s) in a continuum of unknowns (one f value for each s). Apart from a few special cases, in which F has a convenient structure, an exact solution to this problem is computationally intractable.

Instead, we select a function, $\widehat{f}_a$, parameterized by a finite set of coefficients, a, and choose values for a, a^*, to make $F(s;\widehat{f}_a)$ 'small'. Weighted-residual methods compute a^* as the solution to what Reddy (1993, p. 580) refers to as the *weighted-residual form*:

$$\int F(s;\widehat{f}_a)w^i(s)ds = 0, \tag{19}$$

where i ranges from unity to a number which equals the dimension of a. Expression (19) corre-

sponds to a number of equations equal to the number of unknowns in a. The choice of weighting functions in (19) operationalizes the notion of 'small'. For example, if for some i, $w^i = 1$ for all s, then $F(s; \widehat{f}_a)$ small means, among other things, that the average of $F(s; \widehat{f}_a)$, over all possible s, is zero. If for some i, w^i is a Dirac delta function isolating some particular point s, then $F(s; \widehat{f}_a)$ small means it is precisely zero at that point, and so on.

To apply the weighted-residual method, one has to select a family of approximating functions, $\widehat{f}_a$, a set of weighting functions, $w^i(s)$, and strategies for evaluating the integral (19) and any integrals that may go into defining F. The procedures we consider make different choices on these three dimensions. Two general types of $\widehat{f}_a$ functions include *spectral* and *finite element* functions. In the former, each component of a influences $\widehat{f}_a$, over the whole range of s while in the latter, each component of a has influence over only a limited range of s's. We consider three types of weighting functions. In one, the $w^i(s)$'s are related to the basis functions generating $\widehat{f}_a$, in which case the algorithm is an example of the *Galerkin* method. In another, a is chosen so that F is zero at a number of values of s equal to the number of unknown elements in a. In this case, the $w^i(s)$'s are Dirac delta functions, and the algorithm is an example of the *collocation* method. Finally, two numerical procedures are used to evaluate the integrals in (19) and F: quadrature methods and Monte Carlo integration. We now turn to a detailed discussion of several of the algorithms considered.

4. Algorithms for Solving the Model

Some of the algorithms considered are either original to this paper, or are significant modifications to algorithms proposed in the literature. These are reviewed in this section. They are all spectral methods. One approximates the policy function and multiplier function with Chebyshev polynomials. This is an adaptation of the methods advocated in Judd (1992, 1993, 1998). Judd does not consider problems with occasionally binding constraints, which is the focus of our analysis. The other set of spectral methods studied are derived from Marcet's (1988) and Wright and Williams' (1982a,1982b,1984) strategy for approximating conditional expectation functions. In our numerical analysis several finite element methods are considered too. However, we follow

closely the published versions of these algorithms, and so we leave the details to the appendix.

In the first subsection we consider a solution strategy based on parameterizing the policy and multiplier functions. In the second subsection we consider strategies based on approximating conditional expectation functions. To simplify the presentation, we focus on the two-state Markov case, $\theta_t \in \Theta \equiv \{-\sigma, \sigma\}$. Later, we do verify robustness of our numerical results by considering the continuous θ_t case for one model parameterization.

4.1. Parameterizing the Policy and Multiplier Functions

In this subsection, we work with the policy function and Lagrange multiplier characterization of the solution to the model. Consider first the reversible investment version of our model, so that the approximation to h, $\widehat{h}_a$, is identically 0. In this case, we approximate the policy rule as follows:

$$g(k,\theta) \approx \widehat{g}_a(k,\theta) \equiv a_\theta' T\left(\varphi(k)\right), \text{ for } \theta = -\sigma, \sigma, \tag{20}$$

where a_θ is an $N \times 1$ vector of parameters to be solved for, $\theta = -\sigma, \sigma$, and $T(x) = [T_0(x), T_1(x), ..., T_{N-1}(x)]'$. The basis functions for $\widehat{g}_a$, $T_i(x) : [-1,1] \to [-1,1]$, $i = 0, ..., N-1$, are Chebyshev polynomials.[15] Also,

$$\varphi : \left[\underline{k}, \overline{k}\right] \to [-1,1], \ \varphi(k) = 2\frac{k - \underline{k}}{\overline{k} - \underline{k}} - 1. \tag{21}$$

Let $a = [a_\sigma' \ a_{-\sigma}']'$ denote the $2N \times 1$ dimensional vector of parameters for $\widehat{g}_a$.

The $2N$ weighting functions, $w^i(k,\theta)$, are constructed from the basis functions as follows:

$$w^i(k,\theta) = \frac{1}{\left(1 - \varphi(k)^2\right)^{1/2}} \frac{d\widehat{g}_a(k,\theta)}{da_i}, \tag{22}$$

for $i = 1, \ldots, 2N$. It is readily verified from (20) that one of $d\widehat{g}_a(k,\sigma)/da_i$ and $d\widehat{g}_a(k,-\sigma)/da_i$ is zero and the other is a Chebyshev polynomial, for each i.

In the irreversible investment version of the model, we must parameterize the policy and Lagrange multiplier functions so that they respect the Kuhn-Tucker conditions, (4). We impose (and subsequently verify) that the irreversible investment constraint never binds for $\theta = \sigma$.

Thus, we restrict the space of approximating functions for $g(k,\theta)$ as follows:

$$g(k,\sigma) \approx \widehat{g}_a(k,\sigma), \tag{23}$$

$$g(k,-\sigma) \approx \widehat{g}_a(k,-\sigma) \equiv \max\{\widetilde{g}_a(k), \log(1-\delta)+k\}, \text{ for all } k \in [\underline{k}, \overline{k}]. \tag{24}$$

Also,

$$h(k,-\sigma) \approx \widehat{h}_a(k,-\sigma) \equiv \begin{cases} 0, & \widehat{g}_a(k,-\sigma) > \log(1-\delta)+k \\ \max\{\widetilde{h}_a(k), 0\}, & \widehat{g}_a(k,-\sigma) \leq \log(1-\delta)+k \end{cases}. \tag{25}$$

We choose functional forms for $\widehat{g}_a(k,\sigma)$, $\widetilde{g}_a(k)$, and $\widetilde{h}_a(k)$ as follows:

$$\widehat{g}_a(k,\sigma) = a'_\sigma T(\varphi(k)),\ \widetilde{g}_a(k) = a'_{-\sigma} T(\varphi(k)),\ \widetilde{h}_a(k) = b' T(\varphi(k)).$$

Here, T is the $N \times 1$ column vector of Chebyshev polynomials defined after (20), and a_σ, $a_{-\sigma}$, b are $N \times 1$ column vectors of parameters. All elements of a_σ are permitted to be non-zero, while only the first $N_{-\sigma}$ and N_b elements of $a_{-\sigma}$ and b, respectively, can be non-zero. We adopt the restriction $N = N_{-\sigma} + N_b$. Also, let the vector of parameters, a, be composed of the nonzero elements of a_σ, $a_{-\sigma}$, b, so that a has length $2N$. The $2N$ weighting functions are chosen analogously to (22).

The analog of equation (19) is evaluated using $M-$point Gauss-Chebyshev quadrature. To do this, we need the $M \geq N$ grid points, k_j, where

$$k_j = \varphi^{-1}(r_j),\ r_j = \cos\left(\frac{\pi(j-0.5)}{M}\right),\ j = 1, \ldots, M. \tag{26}$$

Here, the r_j's are the M roots of the M^{th} order Chebyshev polynomial, $T_M(x)$. For arbitrary a, the $M-$ point Gauss-Chebyshev quadrature approximation to the weighted residual form of the problem (i.e., the analogue of (19)) is:

$$\begin{aligned} & \int\int R(k,\theta;\widehat{g}_a,\widehat{h}_a) w^i(k,\theta) dk d\theta \\ \approx\ & \frac{\pi}{M}\sum_{j=1}^{M} R(k_j,\sigma;\widehat{g}_a,\widehat{h}_a) w^i(k_j,\sigma) + \frac{\pi}{M}\sum_{j=1}^{M} R(k_j,-\sigma;\widehat{g}_a,\widehat{h}_a) w^i(k_j,-\sigma), \end{aligned}$$

for $i = 1, ..., 2N$. To express this system of equations in matrix terms, we form the $M \times N$ matrix X of rank N:

$$X = \left[\begin{array}{cccc} T(r_1) & T(r_2) & \cdots & T(r_M) \end{array} \right]'. \tag{27}$$

By an orthogonality property of Chebyshev polynomials, the columns of X are orthogonal. Using this notation, the Gauss-Chebyshev quadrature approximation of the weighted residual form is written compactly as follows:

$$X'R(a, \theta) = 0, \ \theta = -\sigma, \sigma, \tag{28}$$

where

$$R(a, \theta) \equiv [R(k_1, \theta; \widehat{g}_a, \widehat{h}_a), R(k_2, \theta; \widehat{g}_a, \widehat{h}_a), \ldots, R(k_M, \theta; \widehat{g}_a, \widehat{h}_a)]'. \tag{29}$$

Expression (28) represents a nonlinear system of $2N$ equations in the $2N$ unknowns, a, which can be solved using widely available software.[16] Below, we refer to this method as Spectral-Galerkin.

For later purposes it is convenient to note that if $M = N$, then X is square and invertible, so the method reduces to setting $R(k_j, \theta; \widehat{g}_a, \widehat{h}_a) = 0$ for $j = 1, ..., M$ and for $\theta = -\sigma, \sigma$. In this case, Spectral-Galerkin reduces to a collocation method.

4.2. Parameterizing the Conditional Expectation

We now discuss methods based on approximating conditional expectations. We distinguish between the type of conditional expectation being approximated and the method used to compute the approximation. We consider two types of conditional expectations, the one that is the focus of Marcet (1988)'s analysis (see e in (8) or (12)) and the one that is the focus in Wright and Williams (1982a, 1982b, 1984) (see v in (14)). We consider two ways of approximating the conditional expectation, one based on the nonlinear regression methods advocated by Marcet (1988) and another that is closely related to the methods advocated by Judd (1992) for approximating policy rules. To simplify the discussion, we focus on methods that approximate the e function and we indicate briefly how the methods must be adjusted to obtain an approximation to v.

For PEAs which approximate e,

$$\int m\left(g(k,\theta),\theta';g,h\right)p_{\theta'}(\theta'\mid\theta)d\theta'\approx\exp[\widehat{e}_a(k,\theta)].$$

Here, $\widehat{e}_a(k,\theta)$ is a function with a finite set of parameters, a. In the reversible investment version of our model economy, $\widehat{h}_a\equiv 0$ and the relation linking the policy function, $\widehat{g}_a$, to $\widehat{e}_a$ can be expressed analytically:

$$\widehat{g}_a(k,\theta)=\log\left\{\exp(\theta+\alpha k)+(1-\delta)\exp(k)-U_c^{-1}\left[\beta\exp\left(\widehat{e}_a(k,\theta)\right)\right]\right\}, \tag{30}$$

where $U_c^{-1}[\cdot]$ is the level of consumption implied by a given value for U_c. In the irreversible investment version of the model, (9)-(10) reduce to

$$\widehat{g}_a(k,\theta)=\log\left[\max\left\{(1-\delta)\exp(k),\exp(\theta+\alpha k)+(1-\delta)\exp(k)-U_c^{-1}\left[\beta\exp\left(\widehat{e}_a(k,\theta)\right)\right]\right\}\right],$$

and

$$\widehat{h}_a(k,\theta)=U_c\left(k,\widehat{g}_a(k,\theta),\theta\right)-\beta\exp[\widehat{e}_a(k,\theta)].$$

We begin by describing a PEA implemented by Marcet (1988), which we refer to as conventional PEA. We then interpret that algorithm as a weighted residual method and use this as a basis for discussing alternative PEAs.

4.2.1. Conventional PEA

In our implementation of conventional PEA, we parameterize the conditional expectation function as follows:

$$\widehat{e}_a(k,\theta)=a_\theta'P(\varphi(k)),\ \text{for}\ \theta=-\sigma,\sigma, \tag{31}$$

where a_θ is the $N\times 1$ vector of parameters to be solved for, and $P(x)=[P_0(x),P_1(x),...,P_{N-1}(x)]'$. The basis functions for $\widehat{e}_a$, $P_i(x):[-1,1]\rightarrow[-1,1]$, $i=0,...,N-1$, are the Legendre polynomials.[17] The function φ is defined in (21), and $a=[a_\sigma'\ a_{-\sigma}']'$ denotes the $2N\times 1$

dimensional vector of parameters for $\widehat{e}_a$.

The conventional PEA applies the following successive approximation method for finding a^*. Before initiating the calculations, simulate a series of length $M+1$, $\{\theta_0, \theta_1, \ldots, \theta_M\}$, using a random number generator. Suppose an initial guess for the $2N$-dimensional parameter vector a is available. A new value, $\tilde{a}$, is computed in two steps:

1. compute $\{k_1, k_2, \ldots, k_{M+1}\}$ recursively from $k_{t+1} = \widehat{g}_a(k_t, \theta_t)$, $t = 0, 1, \ldots M$ using (30) and a given initial value, k_0, and simulate $m_{t+1} = m\left(\widehat{g}_a(k_t, \theta_t), \theta_{t+1}; \widehat{g}_a, \widehat{h}_a\right)$, for $t = 1, ..., M$ using (6),

2. find $\tilde{a}$, the solution to the following nonlinear least-squares regression problem:

$$\tilde{a} = \underset{\bar{a} \in \Re^{2N}}{\operatorname{argmin}} \frac{1}{M} \sum_{t=0}^{M-1} \left[m_{t+1} - \exp(\widehat{e}_{\bar{a}}(k_t, \theta_t))\right]^2 . \tag{32}$$

Let the mapping from a to $\tilde{a}$ defined by the above two steps be denoted by $\tilde{a} = S(a; N, M)$. The conventional PEA seeks a^*, where $a^* - S(a^*; N, M) = 0$, as the limit of the sequence $a, S(a; N, M), S\left[S(a; N, M); N, M\right], \ldots$. As noted by Judd (1998, chapter 13) and Marcet (1988), this algorithm can yield explosive, oscillatory sequences, $a, \tilde{a}, \ldots$, particularly for high values of N. One alternative is to instead iterate on the operator $\widetilde{S}$, where $\widetilde{S}(a; N, M) = (1-\mu)a + \mu S(a; N, M)$, for a small fixed value of μ. A problem with this approach is that it may require time-consuming experimentation with alternative values of μ. In our experience, solving for a^* by applying a quasi-Newton method to the system of equations, $a - S(a; N, M) = 0$, often yields superior results. See Marcet and Marshall (1994) for a discussion of the existence of a^* and of the properties of $\exp[\widehat{e}_{a^*}(k, \theta)]$, $\widehat{g}_{a^*}(k, \theta)$ as M, $N \to \infty$.

Two features of conventional PEA are particularly notable. First, the simulation step which produces the synthetic time series of m_{t+1}'s works with points assigned high probability by $p_{\theta'}$ and $\widehat{g}_a$. Second, conventional PEA involves a nonlinear regression in step 2, which is computationally burdensome.[18] This reflects: (i) the fundamental nonlinearity of the problem, (ii) the large value of M that is required in practice to obtain acceptable accuracy, and (iii) the problems of multicollinearity among regressors that arise in practice for even moderate values

of N (den Haan and Marcet (1990)).

An approximation, $\widehat{v}_a$, to the v function in (14) can be obtained using conventional PEA by implementing a simple adjustment to each of the two steps in the above algorithm. In step 1, the policy and multiplier functions are derived from the parameterized expectation, $\widehat{v}_{\tilde{a}}(k_{t+1}, \theta_t)$, using the mapping defined after (14). In step 2, $\widehat{e}_{\tilde{a}}(k_t, \theta_t)$ is replaced by $\widehat{v}_{\tilde{a}}(k_{t+1}, \theta_t)$. As noted previously, when the θ_t's are iid over time, θ_t can be dropped as an argument in $\widehat{v}_{\tilde{a}}(k_{t+1}, \theta_t)$, reducing the number of parameters to be computed from $2N$ to N.

4.2.2. Conventional PEA as a Weighted Residual Method

To see that the conventional PEA is a particular weighted residual method, note first that for M large and for given a, the first order necessary and sufficient conditions associated with the value of $\tilde{a}$ that solves (32) are:

$$\iiint \left[m\left(\widehat{g}_a(k,\theta), \theta'; \widehat{g}_a, \widehat{h}_a\right) - \exp\left(\widehat{e}_{\tilde{a}}(k,\theta)\right)\right] \exp\left(\widehat{e}_{\tilde{a}}(k,\theta)\right) \frac{d\widehat{e}_{\tilde{a}}(k,\theta)}{da_i} p(k,\theta,\theta';a) dk d\theta d\theta' = 0,$$

for $i = 1, \ldots, 2N$. Here, $p(k, \theta, \theta'; a)$ is the joint density of k, θ, θ', induced by $\widehat{g}_a$ and $p_{\theta'}$. It is readily verified that, for large M, a^* solves the version of (19) with $F = \bar{R}^{pea}$ (for $\bar{R}^{pea}$, see (8)) and weighting functions

$$w^i(k,\theta;a^*) = \frac{p(k,\theta,\theta';a^*)}{p_{\theta'}(\theta' \mid \theta)} \exp(\widehat{e}_{a^*}(k,\theta)) \frac{d\widehat{e}_{a^*}(k,\theta)}{da_i}, \tag{33}$$

for $i = 1, \ldots, 2N$.

In sum, conventional PEA is a weighted residual method that works with the family of approximation functions, $\widehat{g}_a$, defined by (30) and (31); that uses the set of weighting functions given by (33); and that evaluates all integrals by Monte Carlo simulation. The weighting functions emphasize (k, θ, θ') that are assigned high probability by the model. As noted above, this is reflected in step 1 of the conventional PEA, the simulation step.

4.2.3. Alternative Weighted Residual PEAs

Once the conventional PEA is expressed as a weighted residual method, it is clear that there are many other PEAs. Alternative finite parameter functions can be used to parameterize expectations, and there exists a variety of alternative weighting schemes and strategies for evaluating integrals. Here, we discuss one particularly promising class of approaches, which includes the Galerkin and collocation weighted residual methods. We refer to this class as *Chebyshev PEA,* because of its reliance on Chebyshev polynomials as basis functions. Again, the focus of the analysis is on approximating the e function, defined in (12), and we indicate briefly how things must be adjusted when v is the function to be approximated.

The Chebyshev PEAs adopt two modifications on the weighted residual formulation of conventional PEA. First, they work with a slightly modified representation of the residual function, R^{pea}, defined in (12). Substantively, there is no difference between the e function that solves $R^{pea} = 0$ or $\bar{R}^{pea} = 0$. However, we shall see that working with the former allows Chebyshev PEAs to avoid the cumbersome nonlinear regression in step 2 of the conventional PEA. Second, $\widehat{e}_a$ is constructed using Chebyshev polynomials as basis functions. Thus,

$$\widehat{e}_a(k,\theta) = a_\theta' T(\varphi(k)), \text{ for } \theta = -\sigma, \sigma, \tag{34}$$

where a_θ is an $N \times 1$ vector of parameters and $a = [a_{-\sigma}' \; a_\sigma']'$, as before. Also, φ is defined in (21), and $T(\cdot)$ is defined after (20). Advantages of using Chebyshev polynomials as basis functions for $\widehat{e}_a$ are discussed below.

The weighted residual form of the problem is (19) with F replaced by R^{pea} and

$$w^i(k,\theta) = \frac{1}{(1-\varphi(k)^2)^{1/2}} \frac{d\widehat{e}_a(k,\theta)}{da_i},$$

where $\widehat{e}_a$ is defined in (34). As in (28), for arbitrary a, the $M-$ point Gauss-Chebyshev quadrature approximation to (19) is, in matrix form,

$$X' R^{pea}(a,\theta) = 0, \text{ for } \theta = -\sigma, \sigma, \tag{35}$$

where X is an $M \times N$ matrix defined as in (27). Also, the $M \times 1$ vector $R^{pea}(a, \theta)$ is defined analogously to $R(a, \theta)$ in (29).

It is convenient to write (35) in a way that reflects its special structure. Let

$$\hat{e}_a(\theta) = \begin{pmatrix} \hat{e}_a(k_1, \theta) \\ \hat{e}_a(k_2, \theta) \\ \vdots \\ \hat{e}_a(k_M, \theta) \end{pmatrix}, Y_a(\theta) = \begin{bmatrix} \log\left(\int m\left(\widehat{g}_a(k_1, \theta), \theta'; \widehat{g}_a, \widehat{h}_a\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right) \\ \log\left(\int m\left(\widehat{g}_a(k_2, \theta), \theta'; \widehat{g}_a, \widehat{h}_a\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right) \\ \vdots \\ \log\left(\int m\left(\widehat{g}_a(k_M, \theta), \theta'; \widehat{g}_a, \widehat{h}_a\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right) \end{bmatrix}, \text{ for } \theta = -\sigma, \sigma, \tag{36}$$

where $\widehat{g}_a$ is derived from $\hat{e}_a$ using (30). Also,

$$D = (X'X)^{-1}, \; D = \frac{1}{M} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2 \end{bmatrix}.$$

Premultiplying (35) by D and taking into account $\hat{e}_a(\theta) = Xa_\theta$, (35) may be written

$$DX'R^{pea}(a, \theta) = DX'\left[Y_a(\theta) - Xa_\theta\right] = DX'Y_a(\theta) - a(\theta) = 0, \text{ for } \theta = -\sigma, \sigma.$$

Or, stacking this for $\theta = -\sigma$ and σ:

$$S^{pea}(a; N, M) - a = 0, \; S^{pea}(a; N, M) = \begin{bmatrix} DX'Y_a(-\sigma) \\ DX'Y_a(\sigma) \end{bmatrix}. \tag{37}$$

Since (37) is just the individual equations in (35) scaled by non-zero constants, the two systems are equivalent. Thus, finding a^* that solves (35) is equivalent to finding a^* such that $S^{pea}(a^*; N, M) - a^* = 0$.

Consider the following successive approximation method for finding a^*. Before initiating the calculations, compute a fixed set of grid points, k_j, $j = 1, ..., M$, using (26). Suppose a given initial guess for the $2N$-dimensional parameter vector a is available. A new value, $\tilde{a}$, is computed

in two steps:

1. compute the $M \times 1$ vectors, $Y_a(\theta)$, $\theta = -\sigma, \sigma$,

2. find $\tilde{a} = (\tilde{a}'_{-\sigma}, \tilde{a}'_{\sigma})'$, the coefficients in the linear regressions of $Y_a(\theta)$ on the columns of X:

$$\tilde{a}_\theta = (X'X)^{-1} X'Y_a(\theta) = DX'Y_a(\theta),\ \theta = -\sigma, \sigma. \tag{38}$$

If the sequence, a, $S^{pea}(a; N, M), S^{pea}\left[S^{pea}(a; N, M); N, M\right], ...$ converges, then the limit point, a^*, solves (35). We implement an alternative strategy to solve for a^*, by applying a quasi-Newton method to the system of equations, $a - S^{pea}(a; N, M) = 0$. When $M = N$, then X is square and (35) reduces to $R^{pea}(k_j, \theta; \widehat{e}_a) = 0$, $j = 1, ..., N$, $\theta = -\sigma, \sigma$. In this case the algorithm is a collocation method, and we refer to it as PEA collocation. When $M > N$, we refer to this as PEA Galerkin. Each is a special case of Chebyshev PEA.

We can now highlight some of the differences between conventional and Chebyshev PEA. In each case, the heart of the algorithm lies in two steps, a simulation step (step 1) and a regression step (step 2). A distinctive feature of the simulation step under Chebyshev PEA is that a fixed distribution of capital stocks is considered. Later we show that those capital stocks are more widely dispersed relative to the ones considered under conventional PEA. We argue that this feature of Chebyshev PEA permits it to achieve a given amount of accuracy with a smaller value of M than is required for conventional PEA. As for the regression step, it is computationally burdensome under conventional PEA and even breaks down for N large due to multicollinearity reasons. In contrast, under Chebyshev PEA, the regression step is trivial.

To obtain an approximation, $\widehat{v}_a$, to the v function in (14) using Chebyshev PEA simply requires an adjustment to the first step in the above algorithm. Namely, compute $Y_a(\theta)$ as in (36) with $\widehat{g}_a(k_i, \theta)$ replaced by k_i, $i = 1, ..., M$. In addition, the functions $\widehat{g}_a$ and $\widehat{h}_a$ used in constructing $Y_a(\theta)$ are derived from $\widehat{v}_a$ using the mapping defined after (14), with k' evaluated at k_i, $i = 1, ..., M$. As noted above, when θ is iid over time, this modified version of $Y_a(\theta)$ is not a function of θ and so $\widehat{v}_a$ is not either.

4.2.4. The Role of Chebyshev Polynomials in Chebyshev PEA

We see two advantages of using Chebyshev polynomials in Chebyshev PEA. First, the orthogonality property of the columns of X defined after (35) reflects that we construct the grid of k_j's based on the zeros of a Chebyshev polynomial. This is why the linear regression step in (38) is trivial. For example, we have applied the algorithm without difficulty with N as high as 100. In contrast, we had difficulty executing the regression step in conventional PEA (see (32)) for N larger than 5 because of multicollinearity problems.[19] Second, the Chebyshev interpolation theorem (see Judd (1992, 1998)) provides some motivation for selecting the grid of capital stocks based on the roots of a Chebyshev polynomial, at least for PEA collocation. There is some hope that one can establish rigorously that

$$\sup_{(k',\theta)\in\left[\underline{k},\overline{k}\right]\times\Theta} \|\widehat{v}_a - v\| \to 0, \text{ as } N \to \infty,$$

when $M = N$. For further discussion, see Christiano and Fisher (1997).

5. Evaluating the Algorithms

This section evaluates the performance of several algorithms in solving each of seven parameterizations of our model. We consider a broad range of algorithms. This includes the ones described in the previous section. In addition, we consider two finite element methods (FEM) which are described in detail in the appendix. The first of these FEM's, FEM collocation, is a varient of the method implemented by Bizer and Judd (1989), Coleman (1997), Coleman, Gilles and Labadie (1992), Coleman and Liu (1997), and Danthine and Donaldson (1981). It works with the Lagrangian formulation of the planning problem and approximates the policy rule and multiplier function with a piecewise polynomial. As its name suggests, this is a collocation method, in which the weighting functions are Dirac Delta functions. The second FEM, FEM Galerkin, is a varient of the method implemented in McGrattan (1996). It also approximates the policy function with a piecewise polynomial. However, instead of taking the irreversibility constraint into account using a multiplier function, this method makes use of penalty functions. Also, the

weighting function used in this method is the set of basis functions for piecewise polynomials. For convenience, Table 1 contains names and summary descriptions of the various algorithms discussed in this paper.

Implementation of each method requires an initial guess of the solution. For the Lagrange multiplier we use the zero function, and for the policy function we use a standard log-linear approximation, truncated so that gross investment is non-negative.[20] We also obtain a solution to each parameterization using standard dynamic programming methods, and treat this as the 'true' solution for the purpose of evaluating the other algorithms. Details about the dynamic programming method used are reported in the appendix.

In analyzing the properties of various model solutions, we do not examine the computed values of a^*, since these are hard to interpret. Instead, we analyze the implications of a^* for various first and second moment properties of several model variables. We obtain these implications for any particular model solution by simulating a data set of length 100,500, discarding the first 500 observations, and using the rest to compute the first and second moments of interest. In addition to analyzing the second moment implications of the solutions, we also directly examine computed policy functions and the implied Euler equation residuals.

The first subsection that follows discusses the seven model parameterizations that we consider. We then analyze the properties of the PEAs. The final subsection presents a direct comparison of all the algorithms.

5.1. Model Parameterizations Considered in the Analysis

The utility function, U, and Markov transition matrix, $p_{\theta'}$, used in the analysis have the following form:

$$U(c) = \frac{c^{1-\gamma} - 1}{1-\gamma}, \; p_{\theta'} = \begin{bmatrix} \frac{1+\rho}{2} & \frac{1-\rho}{2} \\ \frac{1-\rho}{2} & \frac{1+\rho}{2} \end{bmatrix}.$$

The parameter ρ is the first order autocorrelation of θ and σ is the associated standard deviation. In the benchmark parameterization, labelled model (1) in Table 2, $\beta = 1.03^{-0.25}$, $\gamma = 1.0$, $\alpha = 0.3$, $\delta = 0.02$, $\sigma = 0.23$, $\rho = 0$. The relatively large value of σ was chosen to guarantee that the investment constraint would bind a substantial fraction of times. For the other model

parameterizations, we perturb the benchmark values in the manner indicated in rows (2)-(7) of Table 2. The perturbations were chosen to provide information about the robustness of our results. They include parameterizations with increased curvature in utility (see row (2)) and production ((3) and (4)), and with more persistence and variance ((5) and (6), respectively) in the technology shock. When we increased the curvature in production, we found that σ had to be adjusted simultaneously so that the constraint on investment would continue to bind occasionally. In these cases we adjusted σ so that the constraint binds roughly 20 percent of the time. Row (7) reports a parameterization in which curvature in preferences and technology, and persistence in the technology shock, were increased simultaneously. We also considered a perturbation in which the technology shock is a continuous random variable, and that is discussed below.

Figures 2 and 3 present information about the model solutions for the seven parameterizations. The solid curves graph $I(k,\theta) \equiv g(k,\theta) - \log(1-\delta) - k$, and the price of capital, $q(k,\theta)$, against k for $\theta = \sigma$ and $-\sigma$. In the top two rows of these figures the dashed curves graph $I(k,\theta)$ with the exact $g(k,\theta)$ replaced by its log-linear approximation. Finally, each graph has three vertical lines. The middle one is the nonstochastic steady state value of k and the other two define a symmetric 95 percent confidence interval for k. Several things are worth noting in these figures. First, when $\theta = \sigma$, the non-negativity constraint on investment is never binding. Second, the interval over which it binds when $\theta = -\sigma$ is in most cases in the region of large capital stocks. However, in model (2) it binds for small values of the capital stock.[21] Third, the functions are quite sensitive to model parameterization. In six of the seven parameterizations, investment is decreasing in k when $\theta = -\sigma$ and in the other one it is increasing. Also, the general degree of nonlinearity in the functions varies considerably across parameterizations, although there is always a pronounced kink in the neighborhood where the constraint starts to bind.

5.2. The PEAs

5.2.1. Conventional PEA

Table 3 provides information on the performance of conventional PEA in approximating the conditional expectation, $\exp\left[e(k,\theta)\right]$, that is the focus of Marcet (1988)'s analysis. The results in Marcet and Marshall (1994) indicate that conventional PEA is arbitrarily accurate for sufficiently large M and N. The question that interests us here is how well the algorithm works for the values of M and N used in practice. For the results in Table 3, we set $M = 10,000$. By way of comparison, to solve the growth model, den Haan and Marcet (1990) use $M = 2,500$, den Haan and Marcet (1994) use $M = 29,000$, and den Haan (1995) uses $M = 25,000$. Also, we set $N = 3$. As we shall see, with this value for N and given $M = 10,000$, the benchmark model's implications for the second moment properties of quantities are acceptable.

Recall that a^* obtained by conventional PEA is a function of a random draw of $M+1$ random variables, $\{\theta_0, \theta_1, \ldots, \theta_M\}$. As a result, a^* is itself a random variable. To assess the usefulness of conventional PEA as a solution method, therefore, it is important to consider both bias and Monte Carlo sampling uncertainty in the first and second moment properties implied by approximate solutions obtained with it. To investigate this, we solved each model parameterization $I = 500$ times, each time with an independent random draw, $\{\theta_0, \theta_1, \ldots, \theta_M\}$. When implementing conventional PEA, we always started by trying to use a variant of a quasi-Newton method to solve $a^* - S(a^*; N, M) = 0$. When this method is successful at finding a solution, we found it does so more quickly than does the successive approximation method.

The first three terms in each cluster of four numbers in Table 3 provide information about bias. The unbracketed term is the value of the statistic, s, indicated in the first column implied by the dynamic programming solution. We denote this term by s^{dp}. The term in square brackets, $100(\bar{s} - s^{dp})/s^{dp}$, measures the bias in conventional PEA. Here, $\bar{s}$ is the mean of s across the I conventional PEA solutions. The term in parentheses is the Monte Carlo standard error in the bias statistic in square brackets. The fourth term in each cluster, in angular brackets, measures how much Monte Carlo sampling uncertainty there is in a^*. It reports the coefficient of variation, $100\sigma_s/\bar{s}$, where σ_s is the standard deviation of s across I conventional PEA solutions.

The results in Panel A of Table 3 pertain to various second moment properties of consumption, investment, and output. Here, σ_j, $j = y, c, i$ denote the standard deviation of gross output, consumption and gross investment, respectively, and $\rho(y, j)$, $j = c, i$ denote the correlation of gross output with consumption and gross investment, respectively. The results in Panel B of Table 3 pertain to first and second moment properties of Tobin's q and asset returns.

The results in Panel A indicate that, at least for models (1)-(6), the conventional PEA performs reasonably well. For the most part, bias is not much more than 1 percent. For $\rho(y, c)$, the bias is a little higher in the case of models (1) and (2), where it is about 3.5 percent. The coefficient of variation for these models is also reasonably small, although it is 4.4 percent for $\rho(y, c)$ in model (2). The distortions are somewhat higher for model (7), where the bias in σ_c is 6.3 percent and the associated coefficient of variation is 10.6 percent. Although arguably these last distortions are getting close, none appears to exceed the bounds for acceptability.

According to the information in Panel B, there is greater evidence of distortions in asset prices and returns than in the quantity allocations. For example, even in the benchmark model, the equity premium is understated by roughly 26 percent, and the standard deviation of the equity premium is roughly 20 percent of its average value. Also, the frequency of times that the investment irreversibility constraint is binding (i.e., the frequency of the event, $q < 1$) is understated by 12.5 percent. Still, these distortions do not seem large in an economic sense. The distortions are greater for models (2) - (7). For example, with high risk aversion (model (2)), the standard deviation of the price of capital, q, is overstated by 26.8 percent on average, and its standard deviation across different model solutions is 53 percent of its mean. But, the distortions tend to be largest for statistics involving the rate of return on equity. For example, with model (2) the equity premium is overstated by close to one hundred percent. The performance of conventional PEA deteriorates dramatically for model (7), where quantifying the bias in statistics involving R^e requires scientific notation. To confirm the robustness of this finding, we raised M and N to $50,000$ and 5, respectively, and got very similar results (these results are based on $I = 50$). These are reported in column 2 of Table 4 (column 1 simply reproduces the results from Table 3 for convenience.)

To diagnose the reasons for the poor performance of conventional PEA for model (7), consider

the first four rows of figures in Figure 4. They display the first 20 investment policy rules associated with the $I = 50$ policy rules underlying the calculations in the $N = 5$, $M = 50,000$ column of Table 4. The solid line reports our estimate of the exact investment policy function, $g(k,\theta) - \log(1-\delta) - k$, while the dashed line reports $\widehat{g}_a(k,\theta) - \log(1-\delta) - k$, where $\widehat{g}_a$ is defined in (30), and $\widehat{e}_a(k,\theta)$ was obtained using conventional PEA. Note that in several cases, the approximate investment function obtained using conventional PEA goes to zero for low values of the (log of the) capital stock. When this happens, the estimated price of capital, q, falls below unity, sometimes dropping close to zero.[22] Since q appears in the denominator of the formula for the rate of return on equity (see (18)), when it approaches zero the rate of return on equity rises without bound. Although the zero investment region in Figure 4 occurs with low probability, even very infrequent visits have a dramatic impact on the estimated mean return to equity.

The poor performance of the PEA for financial rates of return reflects that it oversamples the high probability region of the capital stock. One way to see this is to examine the results for 'Modified Conventional PEA' reported in columns 3 and 4 in Table 4. Those are based on a modified version of conventional PEA which samples relatively more heavily from the tails of $[\underline{k}, \overline{k}]$.[23] The modification works by altering step 1 in conventional PEA as follows. We selected five values of the capital stock, $k_1, ..., k_5$, from the interval $[\underline{k}, \overline{k}]$ using the zeros of a fifth order Chebyshev polynomial. Then, corresponding to each (k_i, θ) we drew $5,000$ times from $p_{\theta'}(\theta'|\theta)$ for $i = 1, ..., 5$ and $\theta = -\sigma, \sigma$, respectively. This results in $50,000$ (k, θ, θ') pairs which were used to compute $50,000$ m''s using $m' = m\left(\widehat{g}_a(k,\theta), \theta'; \widehat{g}_a, \widehat{h}_a\right)$. The five capital stocks used by conventional PEA are indicated by the circles in Figure 5B. Note how they are shifted towards the boundaries of the interval $[\underline{k}, \overline{k}]$ relative to a fixed interval grid. For convenience, Figure 5B also displays the density of capital stocks that would result if the 'grid' were obtained using the zeros of a very high order Chebyshev polynomial. The distribution of capital stocks associated with conventional PEA is displayed in Figures 5A, 5C and 5D. These exhibit the model's implications for the unconditional distribution of k, and the distribution of k conditional on $\theta = \sigma$ and $\theta = -\sigma$. The figures confirm that, by comparison with modified conventional PEA, conventional PEA emphasizes capital stocks that are relatively more concentrated in the interior

of $[\underline{k}, \overline{k}]$.[24]

The results in columns 3 and 4 of Table 4 are based on $I = 50$ repetitions of modified conventional PEA. Interestingly, the problems with statistics associated with the rate of return on equity have been dramatically reduced. This reflects that the problems with the investment policy function evident in the top half of Figure 4 have been essentially eliminated (see the bottom half of Figure 4). Bias and coefficient of variation indicates that modified conventional PEA with $N = 5$ and $M = 50,000$ produces a tolerably accurate solution. When M is reduced to $10,000$, bias remains acceptable, but coefficient of variation is now fairly large for statistics related to the rate of return on equity. The improved accuracy that results from increasing dispersion in (k, θ) helps motivate the perturbations in conventional PEA analyzed in the next subsection.

Panel C in Tables 3 and 4 report computation times on a 200 MHz Pentium Pro machine using GAUSS to do the calculations.[25] The times refer to the minimum time needed to solve the model by conventional PEA once. The times for models (1) - (6) are relatively low because our quasi-Newton procedure was successful in these cases. The time for model (7) is higher because the successive approximation method had to be used here. Computation times rise substantially when N and M are increased from 3 to 5, going from roughly six minutes to over one and one-half hours.

5.2.2. Chebyshev PEA

Approximating Marcet's Conditional Expectation Function by PEA Collocation We applied PEA collocation to approximate e in all seven models, and obtained acceptable accuracy with $N = M = 3$ for models (1) to (6). By 'acceptable', we mean that all statistics analyzed in Table 3 and 4 are within 10 percent of their exact values. We only study bias for this method, since Monte Carlo uncertainty is not applicable. Although accuracy for models (1) to (6) was comparable to that obtained by conventional PEA, computation times were drastically lower, closer to one-half second instead of one-half minute or more. To save space, we do not discuss these results and we instead focus on the analysis of model (7).

The last two columns of Table 4 report results using PEA collocation to approximate the

function, e, for model (7). In these columns, we set $N = M = 3$. Figure 6 exhibits the impact of increasing N and M on the Euler errors, $R^{pea}(k, \theta; \hat{e}_a)$ (see (11)). Since the errors are difficult to interpret directly, we convert them into the percent change in consumption needed to make the Euler error zero, holding Tobin's q and the level of investment unchanged.[26]

A notable feature of Figure 6 is that the Euler errors are very small for $N = M = 3$. For example, according to Figure 6, when $\theta = -\sigma$ the $N = M = 3$ rule fails the first order condition by only one, one-hundredth of a percent of consumption. When $\theta = \sigma$ the rule fails by only six, one-thousandths of a percent of current consumption. These are tiny numbers and yet the $N = M = 3$ rule does not produce acceptable accuracy (see Table 4.) To get the desired degree of accuracy, one has to go to $N = M = 5$. We conclude that a researcher interested in financial statistics really must work to make the Euler errors extremely small.

Evidently, the performance of PEA collocation with $N = M = 5$ is comparable or better than that of conventional PEA with $M = 50,000$, $N = 5$, even though the former uses $10,000$ times fewer observations than the latter. This difference is reflected in the amount of computer time required to solve the model. Whereas conventional PEA requires over one and one-half hours to solve the model, PEA collocation requires a little over one-half of a second to get the same degree of accuracy.

Two Further Experiments with Chebyshev PEA

Continuous Exogenous Shock

We considered two other sets of experiments with Chebyshev PEA. In the first we consider a version of model (1) in which the technology shock has a continuous, normal distribution. We did this out of a concern that the experiments in Tables 3 and 4 might be conferring too great an advantage to PEA collocation, over conventional PEA. PEA collocation is in fact compatible with evaluating integrals like those in (8), (12) and (14) by any method whatever, including Monte Carlo methods. However, in most of the experiments with PEA collocation reported in this paper, these integrals were evaluated exactly by fully exploiting the particular two-state distribution assumed for the technology shock. Conventional PEA was not given this advantage. The Monte Carlo method it applies to evaluate these integrals makes no use whatever of the

structure of the shock distribution. To verify that our results are not unduly influenced by this asymmetry of treatment, we also did two experiments with versions of model (1) in which the technology shock has a continuous distribution. The results are reported in Table 5. The column labelled 'benchmark' corresponds to the case in which the shock is iid over time, while the other column corresponds to the case in which the shock has autocorrelation, ρ, equal to 0.95. The integral in (12) was evaluated using $H-$point Gauss-Hermite quadrature integration, with $H = 4$ in each case. The table shows the values of N and M used for conventional PEA and Chebyshev PEA needed to achieve acceptable accuracy, as well as the time needed to execute the computations. The results are consistent with our previous findings. Namely, to get a given degree of accuracy with Chebyshev PEA requires at least an order of magnitude less computation time than does conventional PEA.

Approximating the Wright-Williams Conditional Expectation

For our second set of experiments, we applied PEA collocation to approximating the conditional expectation function, $\exp\left[v(k',\theta)\right]$, emphasized by Wright and Williams (1982a, 1982b, 1984). We address two issues raised in our discussion after (13)-(15): (i) we investigate whether the various potential pathologies discussed there are likely to occur in practice, and (ii) we investigate the relative smoothness of the e and v functions. We work with the benchmark model, model (1), and set $N = M = 5$. Since that model assumes an iid technology shock, θ is not an argument of v. Consequently, application of PEA collocation requires determining the values of only 5 parameters and not the 10 needed to approximate e when $N = 5$.

Our results are displayed in Figure 7. To help assess the accuracy of the calculations, Figure 7A displays the Euler errors, measured in the same units as the errors in Figure 6. The continuous curve indicates the Euler errors over a very fine grid, and the stars indicate the location of the five grid capital grid points used in the calculations.[27] The PEA collocation method forces the Euler errors to be zero at these points. The largest error occurs for k slightly above 4 and is nearly six one hundreths of percent of consumption. The results in Figure 7B allow us to consider issue (i). It displays the initial and final $\hat{v}_a$ functions in the sequence of functions produced in the calculations. As in all other calculations in the paper, the initial function is

the one associated with a zero multiplier function and the log-linearized steady state investment function, truncated so that gross investment is non-negative. The approximate solution was found by initiating the calculations with 10 successive approximation steps, followed by switching to a quasi-Newton procedure. All functions in the sequence generated by this approach are monotonically decreasing, and they rotate smoothly from the initial relatively flat one to the steeper one where the calculations terminated. The fact that these functions are monotonically decreasing is significant, since it establishes that we avoided the various pathologies discussed after (13)-(15).

Figures 7C - 7F allow us to address issue (ii), concerning the relative smoothness of $v(k', \theta)$ versus that of $e(k, \theta) \equiv v\left(g(k, \theta), \theta\right)$. To assess the relative smoothness of these two functions, we compare

$$\begin{aligned} E(k, \theta; a) &\equiv \log\left[\int m\left(\hat{g}_a(k, \theta), \theta'; \hat{g}_a, \hat{h}_a\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right] \text{ and} \\ V(k'; a) &\equiv \log\left[\int m\left(k', \theta'; \hat{g}_a, \hat{h}_a\right) p_{\theta'}(\theta' \mid \theta) d\theta'\right], \end{aligned}$$

where $\hat{g}_a$ and $\hat{h}_a$ are derived from our approximate solution, $\hat{v}_a$. Two sets of observations are relevant in assessing the relative smoothness of E and V. The first can be seen in Figures 7C and 7D, which display gross investment, $\hat{g}_a(k, \theta) - \log(1 - \delta) - k$, for the two lowest values of θ used in the Gauss-Hermite quadrature calculations. Evidently, the slope of $\hat{g}_a(k, \theta)$ undergoes an abrupt jump at the value of the capital stock where the irreversibility constraint becomes binding, a value that is increasing with θ. The second observation is that the function V appears to be quite smooth. This can be seen in Figures 7E and 7F, which plot $dV(k'; a)/dk'$ against k'.[28] These properties of V and $\hat{g}_a$ suggest that V is smoother than E. In particular, they indicate that E must have a kink at the point were the irreversibility constraint becomes binding, since $E(k, \theta; a) = V(\hat{g}_a(k, \theta); a)$.

The presumed lack of smoothness in $E(k, \theta; a)$ is too slight to be visible in a graph of E against k. However, it *is* evident in the graph of the *slope* of E. This can be seen in Figures 7E and 7F, which graph $dE(k, \theta; a)/dk$ against k for the same two values of θ used in Figures 7B and 7C. It is not surprising that the slopes of E and V coincide in the region where the

irreversibility constraint is binding, since the derivative of $\hat{g}_a(k,\theta)$ with respect to k is unity there. Moreover, in light of the negative sign in the slope of V and the fact that the slope of $\hat{g}_a$ jumps when the irreversibility constraint begins to bind, it is also not surprising that the slope of $E(k,\theta;a)$ falls abruptly at this value of k. These observations are consistent with the remarks after equation (17), suggesting that v is a smoother function than e, and therefore easier to approximate numerically.

5.3. Comparing the Algorithms

We now compare conventional PEA and PEA collocation with three other algorithms. The other three include the Spectral-Galerkin method described above, which approximates the policy and multiplier functions with Chebyshev polynomials. In addition, we consider the FEM collocation and FEM Galerkin methods mentioned in the introduction to this section and spelled out in detail in the appendix. We evaluate the algorithms in terms of their ability to achieve a given standard of accuracy and in terms of the seconds of computer time required to do this.

To meet our accuracy standard, a solution must imply a set of values for the 11 statistics studied in Tables 3 and 4 that come within at least 10% of the corresponding exact values obtained by dynamic programming. Our findings are based principally on the results reported in Table 6. With the exception of conventional PEA, results not in parentheses are the minimum, with respect to the algorithm parameters indicated in the table, required for the associated algorithm to achieve the accuracy standard. Entries for conventional PEA simply reproduce the computer times taken from Table 3, which refer to the case $N = 3$ and $M = 10,000$. Finding the minimal computation time for this method was impractical since analysing it requires computing a large number, $I = 500$, solutions for each set of algorithm parameter values. In any case, conventional PEA did not achieve the accuracy standard for models (2), (3), (4) and (7). Thus, for these examples, the reported computer time may be viewed as a lower bound for the true minimal computer time needed to achieve the accuracy standard. Numbers in parentheses in Table 6 report the algorithm parameter values underlying the associated computer times. Although we do not report details on this, the accuracy criterion was hardest to meet for the financial statistics. Most other statistics were within 1% of their exact values.[29]

Several things are worth noting about the results. First, PEA collocation is able to achieve the target degree of accuracy at least ten times faster than any other algorithm. Second, one of the algorithms, Spectral-Galerkin, actually failed to achieve a solution for two of the models. Since all the algorithms were given the same initial values, this failure may reflect a relative weakness of the Spectral-Galerkin approach. A more definitive interpretation of the tendency of this algorithm to crash would require pinpointing the precise reasons for the crashes. We have not done this. Third, conventional PEA fails to meet the targetted level of accuracy in models (2), (3), (4) and (7) (recall Table 3). Table 4 shows that conventional PEA still does not achieve the targetted level of accuracy for model (7), even with $N = 5$, $M = 50,000$. For model (7), conventional PEA uses a particularly large amount of computer time. This reflects that the quasi-Newton algorithm, which worked well for the other models, crashed on this one. As a result, we had to use the slower, successive approximation algorithm described above.

Fourth, neither FEM algorithm consistently dominated the other. FEM collocation is faster that FEM Galerkin in models (2), (3) and (4) and slower in the others. However, when programmer time is also taken into account, we find that FEM Galerkin is considerably slower and more cumbersome to implement than FEM collocation. To see this, recall that FEM Galerkin is a penalty function approach, so that it requires solving a sequence of models, corresponding to an increasing sequence of penalty function parameter values. The computation times reported in Table 6 are the total machine time used for the computations from an initial penalty function parameter value of zero to its final value. For models (1) to (7) the number of penalty function values considered were 20, 979, 23, 68, 23, 72 and 48, respectively. The large number of penalty function values considered for model (2) reflects a need to take relatively small steps in incrementing the penalty function parameter to ensure the algorithm did not crash. Determining the appropriate step size and frequent restarting of the algorithm required considerable programmer intervention. As a result, the times reported in Table 6 greatly understate the actual amount of time needed to implement this algorithm.

In sum, in the examples considered, the PEA collocation algorithm outperforms the others in terms of computer and programmer time. Regarding accuracy, all algorithms except Spectral-Galerkin, outperform conventional PEA in terms of financial statistics. In two examples, the

Spectral-Galerkin algorithm (the one that approximates policy and multiplier functions with polynomial) crashed.

6. Concluding Remarks

In this paper we investigate several algorithms for solving a dynamic model with an occasionaly binding inequality constraint. In an effort to make our results useful to researchers, we apply a variety of algorithms. They are based on a range of different characterizations of the solution to the model. We consider characterizations based on the Lagrangian representation of the solution, the penalty function representation, and representations that characterize the solution in terms of a conditional expectation function. In each case, the algorithm boils down to finding a particular finite parameter function which solves a functional equation. In our analysis of the algorithms, the criteria we consider include computational speed, programming convenience and numerical accuracy. In assessing accuracy, we focussed on model implications for the first and second moment properties of real quantities and financial statistics, including the rate of return on equity. The properties of the latter turned out to be the hardest to approximate well.

One class of algorithms, the PEA, emerged as clearly best in our application. The best known PEA in applied macroeconomics is the one due to Marcet (1988), which we call conventional PEA. Conventional PEA is characterized by the particular conditional expectation that it seeks to approximate, and by the algorithm used to compute the approximation. We document deficiencies in that algorithm, and display an alternative class of approaches, which we call Chebyshev PEA, which eliminates the deficiencies.

We also consider a PEA proposed in Wright and Williams (1982a,1982b,1984), which approximates a different conditional expectation. The conditional expectation approximated by Wright and Williams' PEA is attractive from a computational point of view because it is smooth compared to the function approximated in conventional PEA. At the same time, we show that Wright and Williams' PEA can in principle have other problems. We show that those problems do not actually arise in the analysis of the one sector growth model. Still, they may be of concern in other applications.

In our analysis we were able to evaluate the accuracy of alternative algorithms because of the simplicity of the model economy studied. This allowed us to develop, for comparison purposes, an accurate approximation to the model solution using dynamic programming methods. Of course, in practice this way of assessing accuracy is not available. In a typical application, the best one can do is to attempt to study how successful an algorithm is in driving the relevant functional equation close to zero. We tried to shed light on how close to zero one needs to be to get acceptable accuracy. In our applications, the relevant functional equation corresponds to the Euler error of a planner's first order condition, expressed as a function of the state.[30] We found that to get acceptable accuracy for financial rates of return, especially the rate of return on equity, requires extraordinarily small Euler errors. We measured the error by the percent change in consumption needed to drive that error to zero. We studied one approximate solution in which the maximum Euler error was only 0.012 percent of consumption, and yet there was still an unacceptable 30 percent bias in the mean return on equity.

Although we considered a wide range of parameter values for our model, it bears emphasizing that we have not established that our findings regarding the advantages of various PEAs apply generally.[31] Confidence that results like these hold more generally requires building up experience over time with a variety of applications. We think that further research along these lines would be useful.

A. The Dynamic Programming Algorithm

Throughout our analysis, we approximate the exact solution to our model using a dynamic programming algorithm on a fine grid of capital stocks. We describe the details of our implementation here. It involves first iterating to convergence on a value function and then deriving a decision rule from the converged value function. The mapping that we iterated on is:

$$v_{j+1}(k,\theta) = \max_{k' \in A(k,\theta)} \left\{ u(k,k',\theta) + \beta \left[p_{\theta'}(\sigma \mid \theta) v_j(k',\sigma) + p_{\theta'}(-\sigma \mid \theta) v_j(k',-\sigma) \right] \right\},$$

for $\theta \in \Theta$ and $k' \in \vec{k} = \{k_1, k_2, \ldots, k_M\}$. Also,

$$u(k',k,\theta) = \ln\left[\exp(\theta + \alpha k) + (1-\delta)\exp(k) - \exp(k')\right]$$

and

$$A(k,\theta) = \vec{k} \cap \left\{ k' : \log(1-\delta) + k \leq k' \leq \exp(\theta + \alpha k) + (1-\delta)\exp(k) \right\}.$$

Here, $v_j(\cdot,\sigma)$ and $v_j(\cdot,-\sigma)$ are points in $\Re^M$, $j = 1, 2, \ldots$. Also $v_0(k,\theta) = 0$, for $\theta \in \Theta$ and $k' \in \vec{k}$. The points in $\vec{k}$ are equally spaced with $k_i < k_{i+1}$, $i = 1, 2, \ldots, M-1$, $M = 40,000$. We iterated on the above mapping until reaching a fixed point which was assumed to be achieved when $\mid (v_j - v_{j-1})./v_{j-1} \mid \leq 1 \times 10^{-7}$, here $\mid x \mid$ is the largest element of x in absolute value and $x./y$ represents element by element division of the vectors x and y. Denote the fixed point by v. We then computed the two decision rule vectors $G(\cdot,\sigma)$, $G(\cdot,-\sigma) \in \Re^M$ as follows.

$$G(k,\theta) = \operatorname*{argmax}_{k' \in A(k,\theta)} \left\{ u(k,k',\theta) + \beta \left[p_{\theta'}(\sigma \mid \theta) v(k',\sigma) + p_{\theta'}(-\sigma \mid \theta) v(k',-\sigma) \right] \right\},$$

where $\theta \in \Theta$ and $k' \in \vec{k}$.

The DP second moment properties are based on $G(k,\theta)$ and an imputed multiplier function. The DP version of the multiplier is computed as follows.

$$\lambda(k_i,\theta) = \frac{u_1(k_i, G(k_i,\theta),\theta) - v_1(k_i,\theta)}{1-\delta}, \; i = 1, 2, \ldots, M.$$

Here, u_1 is the derivative of u with respect to its first argument. Also v_1 is our estimate of the derivative of v with respect to its first argument. We obtained this estimate by first fitting, by least squares, a seventh order polynomial to $v(k_i,\theta)$, $i=1,2,\ldots,M$ for $\theta\in\Theta$:

$$v(k_i,\theta)=\beta_0(\theta)+\beta_1(\theta)\varphi(k_i)+\cdots+\beta_7(\theta)[\varphi(k_i)]^7,\ i=1,2,\ldots,M.$$

Here $\varphi:[k_1,k_M]\to[0,1]$. Then,

$$v_1(k_i,\theta)=\left[\beta_1(\theta)+2\beta_2(\theta)\varphi(k_i)+\cdots+7\beta_7(\theta)[\varphi(k_i)]^6\right]\varphi'(k_i),\ i=1,2,\ldots,M.$$

B. Penalty Functions

One of the Finite Element methods that we implement makes use of a penalty function characterization of the model solution. We describe that here. Under this characterization, the solution, g, is the limit of a sequence of solutions, $\{g_n\}$. The n^{th} element in this sequence solves a version of our model in which the irreversible investment constraint, (2), is ignored, and in which the utility function is replaced by $U(c_t)-\frac{\pi_n}{2}\left[\max\{0,(1-\delta)\exp(k_t)-\exp(k_{t+1})\}\right]^2$. Here, $\{\pi_n\}$ is an increasing sequence of positive constants tending toward infinity. The function $g_n:[\underline{k},\overline{k}]\times\Theta\to[\underline{k},\overline{k}]$, satisfies the Euler equation,

$$R^p(k,\theta;g_n,\pi_n)=0,\ \text{for all}\ (k,\theta)\in[\underline{k},\overline{k}]\times\Theta, \tag{39}$$

where

$$\begin{aligned}R^p(k,\theta;g_n,\pi_n)&=U_c(k,g_n(k,\theta),\theta)-\pi_n\max\{0,(1-\delta)\exp(k)-\exp(g_n(k,\theta))\}\\&-\beta\int\{U_c(g_n(k,\theta),g_n(g_n(k,\theta),\theta'),\theta')[f_k(g_n(k,\theta),\theta')+(1-\delta)]\\&-(1-\delta)\pi_n\max[0,(1-\delta)\exp\left(g_n(k,\theta)\right)-\exp\left(g_n\left(g_n(k,\theta),\theta'\right)\right)]\}p_{\theta'}(\theta'\mid\theta)d\theta'=0.\end{aligned}$$

According to Luenberger (1969, section 10.11):

$$g(k,\theta)=\lim_{n\to\infty}g_n(k,\theta),\ \text{and}\ h(k,\theta)=\lim_{n\to\infty}\pi_n\max\{0,(1-\delta)\exp(k)-\exp\left(g_n(k,\theta)\right)\},$$

for each $(k,\theta) \in [\underline{k}, \overline{k}] \times \Theta$. From a computational perspective, an advantage of this characterization is that, for given n, the solution involves only one function, g_n. Moreover, that function need not obey the irreversible investment constraint, (2). A disadvantage of the characterization is that it requires considering many values of n.

C. Finite Element Methods

In this paper we consider only the simplest class of finite element functions, those that are piecewise linear and continuous in k for each fixed θ.[32] We study a collocation (FEM collocation) and Galerkin (FEM Galerkin) procedure for computing the parameters of this function. For FEM collocation, a method advocated by Bizer and Judd (1989), Coleman (1988), Coleman, Gilles, and Labadie (1992) and Danthine and Donaldson (1981), we work with the characterization of the solution in terms of policy and multiplier functions. For FEM Galerkin, a method advocated by McGrattan (1996), we work with the penalty function formulation of the planning problem. We describe the details of our implementation of these algorithms here.

We begin with a description of the policy and multiplier functions relevant to the reversible investment version of the model, so that $\widehat{h}_a \equiv 0$. The $2N \times 1$ vector of parameters of $\widehat{g}_a$, $a = (a'_{-\sigma}, a'_{\sigma})'$, with $a_\theta = (a_{1,\theta}, ..., a_{N,\theta})'$, specify the values of $k' = \widehat{g}_a(k,\theta)$ at each point on a grid of N capital stocks, k_j, $j = 1, ..., N$, for $\theta = -\sigma, \sigma$. Here, $k_1 \geq \underline{k}$, $k_N \leq \overline{k}$ and $k_j < k_{j+1}$, $j = 1, ..., N-1$. We specify that the capital stock grid is composed of equispaced points. Thus, k' corresponding to (k_i, θ) is $a_{i,\theta} = \widehat{g}_a(k_i, \theta)$, for $\theta = -\sigma, \sigma$, $i = 1, 2, \ldots, N$. Policy decisions between points (k_i, θ) are defined by linearly interpolating the decisions at the two nearest such points. Formally,

$$\widehat{g}_a(k,\theta) \equiv a'_\theta L(k), \text{ for } \theta = -\sigma, \sigma. \tag{40}$$

Here, $L(k) = [L_1(k), L_2(k), ..., L_N(k)]'$ is composed of the basis functions for $\widehat{g}_a$:

$$L_i(k) = \begin{cases} \frac{k - k_{i-1}}{k_i - k_{i-1}}, & k_{i-1} \leq k \leq k_i \\ \frac{k_{i+1} - k}{k_{i+1} - k_i}, & k_i \leq k \leq k_{i+1} \\ 0, & \text{elsewhere,} \end{cases}$$

for $i = 2, 3, \ldots N-1$, and

$$L_1(k) = \begin{cases} \frac{k_2 - k}{k_2 - k_1}, & k_1 \leq k \leq k_2 \\ 0, & \text{elsewhere,} \end{cases}, \; L_N(k) = \begin{cases} \frac{k - k_{N-1}}{k_N - k_{N-1}}, & k_{N-1} \leq k \leq k_N \\ 0, & \text{elsewhere.} \end{cases}$$

In the following two subsections, we describe a collocation and a Galerkin procedure for devising a set of $2N$ weighting functions which can be used in conjunction with equation (19) to find a.

C.1. Collocation

Consider first the reversible investment version of the model. FEM-collocation selects values for a so that

$$R(k_i, \theta; \widehat{g}_a, 0) = 0, \tag{41}$$

for $i = 1, 2, \ldots, N$ and $\theta = -\sigma, \sigma$. This is (19), with the weighting functions constructed using suitably chosen Dirac-delta functions. Equation (41) is a nonlinear system of $2N$ equations in the $2N$ unknowns, a. In practice, a method of successive approximation is used to solve (41). In particular, suppose a given initial guess for the parameter vector a is available, and that this is used to define the function, $\widehat{g}_a$, according to (40). A new value, $\tilde{a}$, is computed as follows. For each element of the capital grid k_i and for $\theta = -\sigma, \sigma$, find the $\tilde{a}_{i,\theta}$ that solves

$$\begin{aligned} U_c(k_i, \tilde{a}_{i,\theta}, \theta) = \beta\{&p_{\theta'}(\sigma \mid \theta) U_c(\tilde{a}_{i,\theta}, \widehat{g}_a(\tilde{a}_{i,\theta}, \theta), \sigma)[f_k(\tilde{a}_{i,\theta}; \sigma) + 1 - \delta] \\ &+ p_{\theta'}(-\sigma \mid \theta) U_c(\tilde{a}_{i,\theta}, \widehat{g}_a(\tilde{a}_{i,\theta}, \theta), -\sigma)[f_k(\tilde{a}_{i,\theta}; -\sigma) + 1 - \delta]\}, \end{aligned} \tag{42}$$

for $i = 1, ..., N$. Denote the mapping from a to $\tilde{a}$ by $\tilde{a} = S^F(a; N)$. The method seeks a^*, where $a^* - S^F(a^*; N) = 0$, as the limit of the sequence $a, S^F(a; N), S\left[S^F(a; N); N\right], \ldots$.

Now consider the irreversible investment version of the model. We work with policy and multiplier functions parameterized according to (23)-(25). Accordingly, we choose piecewise linear functions to form $\widehat{g}_a(k, \sigma)$, $\widetilde{g}_a(k)$, and $\widetilde{h}_a(k)$ and select the N-point capital stock grid as in the reversible investment case. The objective is to solve for the coefficients associated with this grid: $a_{i,\sigma}$, $i = 1, 2, \ldots, N$, $a_{i,-\sigma}$, $i = 1, 2, \ldots, N_{-\sigma}$, as before. In addition, we seek b_i,

$i = 1, 2, \ldots, N_b$, where b_i corresponds to $\tilde{h}_a(k_i)$ and $N_b + N_{-\sigma} = N$. Stack the undetermined coefficients in the $2N$ dimensional vector a :

$$a = (a_{1,\sigma}, a_{2,\sigma}, \ldots, a_{N,\sigma}, a_{1,-\sigma}, a_{2,-\sigma}, \ldots, a_{N_{-\sigma},-\sigma}, b_1, b_2, \ldots, b_{N_b})'.$$

We modify the successive approximation algorithm described above as follows. The main step of that algorithm, (42) for $\theta = \sigma$, is replaced by

$$\begin{aligned} U_c(k_i, \tilde{a}_{i,\sigma}, \sigma) = \beta\{p_{\theta'}(\sigma \mid \theta) U_c(\tilde{a}_{i,\sigma}, \widehat{g}_a(\tilde{a}_{i,\sigma}, \sigma), \sigma)[f_k(\tilde{a}_{i,\sigma}, \sigma) + 1 - \delta] \\ + p_{\theta'}(-\sigma \mid \theta)(U_c(\tilde{a}_{i,\sigma}, \widehat{g}_a(\tilde{a}_{i,\sigma}, -\sigma), -\sigma)[f_k(\tilde{a}_{i,\sigma}, -\sigma) + 1 - \delta] - \widehat{h}_a(\tilde{a}_{i,\sigma}, -\sigma)(1 - \delta))\}. \end{aligned} \tag{43}$$

Equation (42) for $\theta = -\sigma$ is replaced by:

$$\begin{aligned} U_c(k_i, \tilde{a}_{i,-\sigma}, -\sigma) - b_i = \beta\{p_{\theta'}(\sigma \mid \theta) U_c(\tilde{a}_{i,-\sigma}, \widehat{g}_a(\tilde{a}_{i,-\sigma}, \sigma), \sigma)[f_k(\tilde{a}_{i,-\sigma}, \sigma) + 1 - \delta] \\ + p_{\theta'}(-\sigma \mid \theta)(U_c(\tilde{a}_{i,-\sigma}, \widehat{g}_a(\tilde{a}_{i,-\sigma}, -\sigma), -\sigma)[f_k(\tilde{a}_{i,-\sigma}, -\sigma) + 1 - \delta] - \widehat{h}_a(\tilde{a}_{i,-\sigma}, -\sigma)(1 - \delta))\}. \end{aligned} \tag{44}$$

(Recall, we impose - and later verify - that the irreversibility constraint never binds in the high shock state.) For each i, equation (43) is solved by choice of $\tilde{a}_{i,\sigma}$, as before. Equation (44) is first solved by choice of $\tilde{a}_{i,-\sigma}$ with $b_i = 0$. If $\tilde{a}_{i,-\sigma} > [\log(1 - \delta) + k_i]$ then we proceed to the next value of i in the sequence $i = 1, 2, \ldots, N$. Otherwise, $\tilde{a}_{i,-\sigma}$ is set equal to $[\log(1 - \delta) + k_i]$ and (44) is solved by choice of b_i. In this way, (43) and (44) define a mapping from a to $\tilde{a}$, as before. The method iterates on this mapping until convergence.

C.2. Galerkin

Consider the reversible investment case first, so that $\widehat{h}_a \equiv 0$. In our example, the method works to select the value of a that solves the analog of (19) with $w^i(k, \theta) = d\widehat{g}_a(k, \theta)/da_i$, $i = 1, ..., 2N$. Taking into account the region over which L_i is zero, equation (19) is:

$$\begin{aligned} \int_{k_{i-1}}^{k_{i+1}} R(k, \theta, \widehat{g}_a, 0) L_i(k) dk &= 0 \text{ for } i = 2, \ldots, N-1 \\ \int_{k_i}^{k_{i+1}} R(k, \theta, \widehat{g}_a, 0) L_i(k) dk &= 0 \text{ for } i = 1 \end{aligned} \tag{45}$$

$$\int_{k_{i-1}}^{k_i} R(k,\theta,\widehat{g}_a,0)L_i(k)dk \;\; = \;\; 0 \text{ for } i = N,$$

for $\theta = -\sigma, \sigma$. Here, R is defined in (3). We approximated each integral using M-point Gauss-Legendre quadrature integration (see Press, Flannery, Teukolsky, and Vetterling (1992, pp. 140-153)). The approximations yield a $2N-$equation system that can be used to solve for the $2N$ unknowns, a, as in (28) or (35). We used a nonlinear equation solver to actually do the calculations.[33]

Now consider the irreversible investment case. We work with the characterization of the solution based on penalty functions. The penalty function method solves a sequence of systems of equations like (45), in which $R(k,\theta,\widehat{g}_a,0)$ is replaced by $R^p(k,\theta;\widehat{g}_a,\pi_n)$, defined in (39). We considered an increasing sequence of penalty function weights, π_1, π_2, ..., and stopped when the maximum violation of the irreversible investment constraint, (2), over (k_i,θ), $i = 1,...,N$, and $\theta = -\sigma,\sigma$ is smaller than some prespecified tolerance. Denote by π^* the value of the penalty parameter when the algorithm stopped, and let a^* denote the associated value of the parameter vector, a. Then, following Luenberger (1969, Theorem 2, p. 307), our approximation to $h(k,\theta)$ is given by:

$$\widehat{h}_{a^*}(k,\theta) = \pi^* \max\{0,(1-\delta)\exp(k) - \exp\left(\widehat{g}_{a^*}(k,\theta)\right)\}.$$

Notes

[1]See, for example, Aiyagari (1993), Aiyagari and Gertler (1991), Coleman and Liu (1997), den Haan (1996a), Heaton and Lucas (1992, 1996), Huggett (1993), Kiyotaki and Moore (1997), Marcet and Ketterer (1989), Marcet and Marimon (1992), Marcet and Singleton (1990), Telmer (1993), and McCurdy and Ricketts (1995).

[2]For an example of the former, see Atkeson and Kehoe (1993), Boldrin, Christiano and Fisher (1995) and Coleman (1997). Examples of the latter include Gustafson (1958), Aiyagari, Eckstein and Eichenbaum (1980), Wright and Williams (1982a, 1982b, 1984), Miranda and Helmberger (1988), Christiano and Fitzgerald (1991) and Kahn (1992).

[3]For example, solving the heterogeneous agent models of Aiyagari (1993), Aiyagari and Gertler (1991) and Huggett (1993) requires repeatedly solving a partial equilibrium asset accumulation problem for an individual agent, for different values of a particular market price. A solution to the general equilibrium problem is obtained once a value for the market price is found which implies a solution to the partial equilibrium problem that satisfies a certain market clearing condition. The partial equilibrium model solved in these examples is similar to the growth model we work with in this paper.

[4]There are several algorithms that we were not able to include in our anlaysis. One is an interesting one due to Paul Gomme (1997), based on ideas from adaptive learning. Others include the algorithms due to Greenwood, McDonald and Zang (1995), Hartley (1996), and Deborah Lucas (1994). For an analysis of many of the algorithms discussed here as applied to the growth model with reversible investment, see the collection of papers summarized in Taylor and Uhlig (1990). For other useful discussions of solution methods, see Rust (1996) and Santos (1997).

[5]See Judd (1992, 1993, 1998) for recent discussions of this example.

[6]In many models, we expect that nonlinearities in the functions being approximated are greatest in the tail areas. This suggests that oversampling these areas may require increasing the

degree of nonlinearity in the approximating functions beyond what is customary in conventional practice.

[7]For other applications in which this method has been applied successfully, see den Haan (1996b, 1997), den Haan, Ramey and Watson (1997) and Stockman (1997).

[8]The modification must take into account that under (2) the constraint set for capital does not satisfy monotonicity.

[9]Sufficient conditions for a solution include not just the Kuhn-Tucker and Euler equations, but also a transversality condition. A sufficient condition for the latter is that a given candidate solution imply a bounded ergodic set for capital. This result is what we use in practice to verify that our candidate approximate solutions satisfy the transversality condition.

[10]To establish that $\mathcal{P}(v)$ is decreasing in its first argument if v is, (16) indicates it is sufficient to verify that m is decreasing in its first argument whenever v is. Accordingly, consider a given $v(k,\theta)$ that is decreasing in k for $(k,\theta) \in [\underline{k},\overline{k}] \times \Theta$. Fix $\theta \in \Theta$ and consider first values of k interior to the set of points where the irreversibility constraint fails to bind. From the relation, $U_c\left(k, g(k,\theta),\theta\right) = \beta \exp\left[v(g(k,\theta),\theta)\right]$, it is easily verified that $U_c\left(k, g(k,\theta),\theta\right)$ is increasing in k. But, $m(k,\theta;g,h) = U_c\left(k, g(k,\theta),\theta\right)\left[f_k(k,\theta) + 1 - \delta\right]$. The result that m is decreasing in its first argument follows from the fact that U_c and f_k are. Now suppose k lies in the interior of the set where the irreversibility constraint binds, so that $g(k,\theta) = \log(1-\delta) + k$. Then, substituting (15) into (6), we get $m(k,\theta;g,h) = U_c\left(k, g(k,\theta),\theta\right) f_k(k,\theta) + (1-\delta)\beta \exp\left[v(k,\theta)\right]$. That m is decreasing in k follows from the readily verified facts that U_c, v and f_k are.

[11]We implicitly assume that the value(s) of k where $g(k,\theta)$ has a kink varies nontrivially with θ. For an illustration of the proposition that the second type of kink has a relatively small impact on e, see Figure 2 in den Haan (1997). The solid line in that figure graphs the analog of our e function for a wealth accumulation problem in which the exogenous shock (the analog of our θ') can take on two values. The curve in den Haan's figure has two kinks. The more pronounced one is an example of our first type of kink, whereas the smaller one is an example of our second type.

[12]The following example illustrates some of the points in the preceding two paragraphs. Suppose $m(k',\theta') = \max(k',\theta')$, for $k',\theta' \in [\underline{k},\overline{k}]$. Let $f(k') \equiv \int_{\underline{k}}^{\overline{k}} m(k',\theta')d\theta' = k'(k'-\underline{k})+0.5(\overline{k}-k')^2$. The function f, which is the analog of v, is clearly differentiable in k' even though $m(k',\theta')$ is not. Now, suppose $g(k)$ is some non-differentiable function. Then the composite function, $\tilde{f}(g) = f(g(k))$, which is the analog of e, is not differentiable either. Thus, f is smoother than $\tilde{f}$.

[13]When the technology shock is *iid*, then v is not even a function of θ. This is a great simplification from a computational perspective.

[14]See Sargent (1980) and Christiano and Fisher (1998) for a more detailed analysis of Tobin's q in a general equilibrium environment like ours.

[15]The Chebyshev polynomials are defined as follows: $T_0(x) \equiv 1$, $T_1(x) = x$, and $T_i(x) = 2xT_{i-1}(x) - T_{i-2}(x)$, for $i \geq 2$.

[16]Here and throughout the paper, we apply the version of the quasi-Newton algorithm with Broyden's secant update method implemented in the GAUSS procedure NLSYS.

[17]The i^{th} polynomial is $P_i(x) = 1 + \alpha_1^i x + \ldots + \alpha_i^i x^i$, with the α's defined by the requirement $P_0(x) \equiv 1$ and $\int_{-1}^{1} P_i(x)P_j(x)dx = 0$ for $j = 0,\ldots,i-1$ and $i \geq 1$. These polynomials were chosen to help mitigate possible computational problems arising from multicollinearity in step 2 of the conventional PEA, which is discussed below. We have not investigated whether computational results are sensitive to the choice of polynomial. Mathematically, there is no sensitivity.

[18]See Marshall (1992) for a discussion of strategies for easing the burden of this step.

[19]We did not do the experiments necessary to determine whether the multicollinearity problems we encountered would have been less or greater if we had used ordinary polynomials instead of Legendre polynomials.

[20]Details of the log-linear approximation procedure we used are described in Christiano (1991, Appendix). We initiate the PEA calculations by using the multiplier and policy functions just

described the first time the simulation step (step 1) is executed.

[21]It is not apparent in Figure 2, but the region in which the constraint binds in model (2) is strictly interior to the ergodic set for capital.

[22]See Figure 5 in Christiano and Fisher (1997), for graphs of q.

[23]Our modified conventional PEA is similar to what Marcet and Marimon (1992) refer to as *PEA with exogenous oversampling.* They argue that by increasing the dispersion in capital relative to conventional PEA, one gets a more accurate estimate of the far-from-steady-state properties of a model. Our analysis suggests that this observation may even apply when the objects of interest are properties of the steady state distribution implied by the model.

[24]Note that in general the distribution of capital stocks used with conventional PEA does not have to correspond to the distribution implied by the true model solution. Figures 5A, 5C and 5D have been constructed using the true model solution.

[25]The simulation portion of conventional PEA was coded as a FORTRAN subroutine and imported to GAUSS using the GAUSS foreign language interface. This was to combat the well-known deficiency of GAUSS with respect to long do-loops.

[26]Let c denote the level of consumption in the approximate solution and let $\tilde{c}$ denote the level of consumption needed to set the Euler error to zero without changing either the level of investment or Tobin's q. We have $c^{-\gamma} - \hat{h}_a(k,\theta) = qc^{-\gamma} = \beta \exp\left(\hat{e}_a(k,\theta)\right)$, and $\tilde{c}$ is defined by the relation, $\tilde{c}^{-\gamma} q = \beta \left[\exp\left(\hat{e}_a(k,\theta) - R^{pea}(k,\theta;\hat{e}_a)\right)\right]$. Dividing and rearranging, we get our consumption-based measure of the Euler error: $100\left(\tilde{c}/c - 1\right) = 100\left[\exp\left(R^{pea}(k,\theta;\hat{e}_a)/\gamma\right) - 1\right]$.

[27]The interval width for the fine grid used in Figure 7 is 0.005.

[28]In this and the next paragraph, $df(x)/dx$ means $[f(x) - f(x-\varepsilon)]/\varepsilon$ for $\varepsilon = 0.005$.

[29]With regard to conventional PEA, these results are consistent with den Haan (1995)'s findings. He reports that he needed very high values of M to get accurate solutions for rates of return with conventional PEA.

[30]Santos (1998) takes some steps in the direction of developing formal methods for using Euler equation errors to assess the accuracy of an approximation. He does not consider the case of occasionally binding constraints studied here. In addition, in assessing accuracy we focus on implications for second moment properties, while Santos focuses on implications for policy rules.

[31]An exception, noted in the introduction, is that we do show that the linearity and orthogonality properties of Chebyshev PEA apply in arbitrary dimensions.

[32]Reddy (1993) describes systematic procedures for expanding the space of finite element functions to include more than one dimension, and piecewise polynomials of order higher than one.

[33]To see exactly how we do this, write the typical integral in (45) as $\int_a^b R(k,\theta,\widehat{g}_a)M_i(k)dk$ where, for example, $a = k_{i-1}$, $b = k_{i+1}$ when $i = 2,...,N-1$. The Gauss-Legendre M-point quadrature approximation to this integral is written (*) $\sum_{j=1}^{M} R(k_j,\theta,\widehat{g}_a)M_i(k_j)v_j$, where the algorithm for computing the v_j's is provided in (Press, et al. (1992)). To compute the k_j's, we first solve for r_j, $j = 1,...,M$, the M zeros of the M^{th} order Legendre polynomial, $P_M(x)$, discussed after (31). The r_j's and v_j's depend only upon the parameter, M. Then, $k_j = [r_j(b-a)+(a+b)]/2$, $j = 1,...,M$. In this way, we get N equations like (*). These can be represented in matrix form, as in (28) or (35), where the analog of X is composed of the basis functions of $\widehat{g}_a$, the v_j's and the k_j's. In contrast with the case of Spectral-Galerkin and Chebyshev-PEA, the columns of X are not orthogonal.

References

Aiyagari, S. R., 1993, Frictions in Asset Pricing and Macroeconomics, *European Economic Review* 38, 932-39.

Aiyagari, S. R., Z. Eckstein, and M.Eichenbaum, 1980, Rational Expectations, Inventories and Price Fluctuations, Yale University Economic Growth Center Discussion Paper 363, October.

Aiyagari, S. R. and M. Gertler, 1991, Asset Returns With Transactions Costs and Uninsured Individual Risk, *Journal of Monetary Economics* 27, 311-331.

Atkeson, A. and P. Kehoe, 1993, Industry Evolution and Transition: The Role of Information Capital, manuscript.

Bizer, D.S. and K.L. Judd, 1989, Uncertainty and Taxation, *American Economic Review79, 331-336.*

Boldrin, M., L.J. Christiano, and J.D.M. Fisher, 1995, Asset Pricing Lessons for Modeling Business Cycles, National Bureau of Economic Research Working Paper No. 5262.

Christiano, L.J, 1991, Modeling the Liquidity Effect of a Monetary Shock, Federal Reserve Bank of Minneapolis *Quarterly Review* 15, 3–34.

Christiano, L.J., and J.D.M. Fisher, 1994, Algorithms for Solving Dynamic Models with Occasionally Binding Constraints, Federal Reserve Bank of Minneapolis Staff Report 171.

Christiano, L.J., and J.D.M. Fisher, 1997, Algorithms for Solving Dynamic Models with Occasionally Binding Constraints, National Bureau of Economic Research Technical Working Paper 218.

Christiano, L.J., and J.D.M.Fisher, 1998, Stock Market and Investment Good Prices: Implications for Macroeconomics, Federal Reserve Bank of Chicago working paper.

Christiano, L.J. and T. Fitzgerald, 1991, The Magnitude of the Speculative Motive for Holding Inventories in a Business Cycle Model, Federal Reserve Bank of Minneapolis Working Paper 415.

Coleman, W.J., 1988, Money, Interest and Capital in a Cash-in-Advance Economy, International Finance Discussion Paper No. 323, Board of Governors of the Federal Reserve System.

Coleman, W.J., 1997, The Behavior of Interest Rates In a General Equilibrium, Multisector Model with Irreversible Investment, *Macroeconomic Dynamics* 1 206-227.

Coleman, W.J., and M. Liu, 1997, Incomplete Markets and Inequality Constraints: Some Theoretical, Computational, and Empirical Issues, Duke University Fuqua School of Business manuscript.

Coleman, W.J., C. Gilles, and P. Labadie, 1992, The Liquidity Premium in Average Interest Rates, *Journal of Monetary Economics* 30, 449-465.

Danthine, J-P. and J.B. Donaldson, 1981. Stochastic Properties of Fast vs. Slow Growing Economies, *Econometrica* 49, 1007-1033.

den Haan, W, 1995, The Term Structure of Interest Rates in Real and Monetary Economies, *Journal of Economic Dynamics and Control* 19, 909-40.

den Haan, W., 1996a, Heterogeneity, Aggregate Uncertainty and the Short Term Interest Rate: A Case Study of Two Solution Techniques, *Journal of Business and Economic Statistics* 14, 399-411.

den Haan, W., 1996b, Understanding Equilibrium Models with a Small and a Large Number of Agents, National Bureau of Economic Research Working Paper 5792.

den Haan, W., 1997, Solving Dynamic Models With Aggregate Shocks and Heterogeneous Agents, *Macroeconomic Dynamics* 1, 355-386.

den Haan, W. and A. Marcet, 1990, Solving the Stochastic Growth Model by Parameterizing Expectations, *Journal of Business and Economic Statistics* 8, 31-34.

den Haan, W. and A. Marcet, 1994, Accuracy in Simulation, *Review of Economic Studies* 61, 3-18.

den Haan, W., G. Ramey, and J. Watson, 1997, Job Destruction and Propagation of Shocks. University of California at San Diego manuscript.

Gomme, P., 1997, Evolutionary Programming as a Solution Technique for the Bellman Equation, Federal Reserve Bank of Cleveland manuscript.

Greenwood, J., G.M., MacDonald, G-J Zhang, 1996, The Cyclical Behavior of Job Creation and Job Destruction: A Sectoral Model. *Economic Theory* 7, 95-112.

Gustafson, R.L., 1958, Carryover Levels for Grains: A Method for Determining Amounts that are Optimal Under Specified Conditions, Technical Bulletin No. 1178, US Department of Agriculture, Washington, D. C.

Hartley, P., 1996, Value-Function Approximation in the Presence of Uncertainty and Inequality Constraints: An Application to the Demand for Credit Cards, *Journal of Economic Dynamics and Control* 20, 63-92.

Heaton, J. and D. Lucas, 1992, The Effects of Incomplete Insurance Markets and Trading Costs in a Consumption-Based Asset Pricing Model, *Journal of Economic Dynamics and Control* 16, 601-620.

Heaton, J. and D. Lucas, 1996, Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing, *Journal of Political Economy* 104, 443-87.

Huggett, M., 1993, The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies, *Journal of Economic Dynamics and Control* 17, 953-969.

Judd, K., 1992, Projection Methods for Solving Aggregate Growth Models, *Journal of Economic Theory* 58, 410-452.

Judd, K., 1994, Comments on Marcet, Rust and Pakes, in C. Sims, ed., Advances in Econometrics, Sixth World Congress, Vol 2 (Cambridge University Press, Cambridge).

Judd, K., 1998, *Numerical Methods in Economics,* (MIT Press, Cambridge), forthcoming.

Kahn, J., 1992, Why is Production More Volatile than Sales? Theory and Evidence on the Stock-out Avoidance Motive for Inventory Holding, *Quarterly Journal of Economics* 107, 481-510.

Kiyotaki, N. and J. Moore, 1997, Credit Cycles, *Journal of Political Economy* 105, 211-48.

Lucas, D., 1994, Asset Pricing with Undiversifiable Income Risk and Short Sales Constraints: Deepening the Equity Premium Puzzle, *Journal of Monetary Economics* 34, 325-41

Luenberger, D., 1969, *Optimization by Vector Space Methods* (John Wiley & Sons, New York).

Marcet, A., 1988, Solving Nonlinear Stochastic Growth Models by Parametrizing Expectations, Carnegie-Mellon University manuscript.

Marcet, A. and J. Ketterer, 1989, Introducing Derivative Securities: A General Equilibrium Approach, Carnegie-Mellon University manuscipt.

Marcet, A. and R. Marimon, 1992, Communication, Committment and Growth, *Journal of Economic Theory* 58, 219-249.

Marcet, A. and D. Marshall, 1994, Solving Non-linear Rational Expectations Models by Parameterized Expectations: Convergence to Stationary Solutions, Federal Reserve Bank of Chicago Working Paper 94-20.

Marcet, A. and K. Singleton, 1990, Equilibrium Asset Prices and Savings of Heterogeneous Agents In the Presence of Portfolio Constraints, Carnegie-Mellon manuscript.

Marshall, D., 1992, Inflation and Asset Returns in a Monetary Economy, *Journal of Finance* 47, 1315-42.

McCurdy, T. and N. Ricketts, 1995, An international economy with country-specific money and productivity growth processes. *Canadian Journal of Economics* 28, 141-62.

McGrattan, E., 1996, Solving the Stochastic Growth Model with a Finite Element Method, *Journal of Economic Dynamics and Control* 20, 19-42.

Miranda, M.J., 1985, Analysis of Rational Expectations Models for Storable Commodities Under Government Regulation, University of Wisconsin-Madison Doctoral Dissertation.

Miranda, M.J., and P.G. Helmberger, 1988, The Effects of Commodity Price Stabilization Programs, *American Economic Review* 78, 46-58.

Press, W.H., S. Teukolsky, W. Vetterling, and B. Flannery, 1992, *Numerical Recipes in Fortran* (Cambridge University Press, New York).

Reddy, J.N., 1993, *An Introduction to the Finite Element Method* (McGraw-Hill,New York).

Rust, J., 1996, Numerical Dynamic Programming in Economics, in H. Amman, D. Kendrick and J. Rust, eds., *Handbook of Computational Economics* (North Holland, Amsterdam) 619-729.

Santos, M., 1997, Numerical Solution of Stochastic Growth Models, University of Minnesota manuscript.

Santos, M., 1998, An Accuracy Test Using the Euler Equation Residuals, University of Minnesota manuscript.

Sargent, T., 1980, Tobin's q and the Rate of Investment in General Equilibrium, *Carnegie-Rochester Conference Series on Public Policy* 12, 107-54.

Stockman, D.R., 1997, Balanced-Budget Rules: Welfare Consequences, Optimal Policies, and Theoretical Implications. University of Chicago Doctoral Dissertation.

Stokey, N. and R.E. Lucas, Jr., with E. Prescott, 1989, *Recursive Methods in Economic Dynamics* (Harvard University Press, Cambridge, MA).

Taylor, J., and H. Uhlig, 1990, Solving Nonlinear Stochastic Growth Models: A Comparision of Alternative Solution Methods, *Journal of Economics and Business* 8, 1-17.

Telmer, C., 1993, Asset Pricing Puzzles and Incomplete Markets, *Journal of Finance* 48, 1803-32.

Wright, B.D., and J.C. Williams, 1982a, The Economic Role of Commodity Storage, *Economic Journal* 92, 596-614.

Wright, B.D., and J.C. Williams, 1982b, The Roles of Public and Private Storage in Managing Oil Import Disruptions, *Bell Journal of Economics* 13, 341-353.

Wright, B.D., and J.C. Williams, 1984, The Welfare Effects of the Introduction of Storage, *Quarterly Journal of Economics* 99, 169-182.

Table 1. Summary of the computational strategies considered

Computational Strategy[(i)]	Object Approximated	Residual Weighting Scheme	Evaluation of Integrals
	Spectral Methods[(ii)]		
conventional PEA	Marcet conditional expectation	model-implied density for capital and technology	Monte Carlo
modified conventional PEA	Marcet conditional expectation	exogenous density for capital and technology	Monte Carlo
Chebyshev PEA	Marcet and Wright-Williams conditional expectation	dirac delta functions (if collocation) Galerkin (if Galerkin)	quadrature
PEA collocation	Marcet and Wright-Williams conditional expectation	dirac delta functions	quadrature
PEA Galerkin	Marcet conditional expectation	Galerkin	quadrature
Spectral-Galerkin	policy and multiplier functions	Galerkin	quadrature
	Finite Element Methods[(iii)]		
FEM collocation	policy and multiplier functions	dirac delta functions	quadrature
FEM Galerkin	policy function	Galerkin	quadrature

Notes to table 1.

(i) These names are intended as a convenient shorthand only. For example, technically PEA Galerkin is a Spectral Galerkin method too.

(ii) We used polynomials.

(iii) We used piecewise linear functions.

Table 2. Parameterizations considered

	Parameter Values					
Model	β	γ	α	δ	σ	ρ
(1)	$1.03^{1/4}$	1	0.3	0.02	0.23	0
(2)	$1.03^{1/4}$	10	0.3	0.02	0.23	0
(3)	$1.03^{1/4}$	1	0.05	0.02	0.0382	0
(4)	$1.03^{1/4}$	1	0.3	0.5	0.675	0
(5)	$1.03^{1/4}$	1	0.3	0.02	0.23	0.95
(6)	$1.03^{1/4}$	1	0.3	0.02	0.40	0
(7)	$1.03^{1/4}$	10	0.1	0.02	0.23	0.95

Table 3. Bias and Monte Carlo variation in conventional PEA

Statistic	Parameterizations (1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A – Quantities				
σ_y	66.0	67.9	3.95	68.8	84.8	125.0	34.2
	[−0.2]	[−0.6]	[−0.04]	[−0.3]	[0.1]	[−0.3]	[−0.2]
	(0.004)	(0.02)	(0.001)	(0.01)	(0.02)	(0.003)	(0.04)
	⟨0.1⟩	⟨0.5⟩	⟨0.01⟩	⟨0.1⟩	⟨0.5⟩	⟨0.1⟩	⟨0.9⟩
σ_c	10.2	7.50	1.01	34.3	49.7	45.4	12.4
	[−1.1]	[1.1]	[−0.2]	[−0.9]	[−0.2]	[−1.0]	[6.3]
	(0.03)	(0.1)	(0.04)	(0.01)	(0.02)	(0.01)	(0.5)
	⟨0.7⟩	⟨1.6⟩	⟨0.9⟩	⟨0.3⟩	⟨0.⟩	⟨0.3⟩	⟨10.6⟩
σ_i	61.9	65.9	3.42	36.1	43.8	80.8	23.1
	[0.04]	[−0.8]	[0.3]	[−0.1]	[0.3]	[0.2]	[−3.8]
	(0.01)	(0.03)	(0.01)	(0.02)	(0.1)	(0.01)	(0.3)
	⟨0.2⟩	⟨0.6⟩	⟨0.2⟩	⟨0.4⟩	⟨1.4⟩	⟨0.3⟩	⟨7.3⟩
$\rho(y,c)$	0.47	0.32	0.61	0.98	0.92	0.98	0.94
	[−3.0]	[3.5]	[−0.3]	[−0.3]	[−0.1]	[0.2]	[−0.2]
	(0.05)	(0.2)	(0.02)	(0.001)	(0.01)	(0.002)	(0.1)
	⟨1.1⟩	⟨4.4⟩	⟨0.5⟩	⟨0.03⟩	⟨0.3⟩	⟨0.1⟩	⟨1.1⟩
$\rho(y,i)$	0.99	0.99	0.97	0.98	0.90	0.99	0.98
	[−0.1]	[0.40]	[0.3]	[0.01]	[−0.5]	[0.5]	[−0.2]
	(0.001)	(0.001)	(0.002)	(0.002)	(0.02)	(0.001)	(0.03)
	⟨0.01⟩	⟨0.02⟩	⟨0.1⟩	⟨0.04⟩	⟨0.4⟩	⟨0.02⟩	⟨0.7⟩
			Panel B – Asset Prices and Returns				
ER^e	3.20	3.08	3.01	309.5	2.94	59.8	1.44
	[−0.1]	[25.1]	[0.4]	[2.0]	[−0.2]	[−1.7]	[1.6 e 8]
	(0.05)	(1.9)	(0.02)	(0.1)	(0.1)	(0.04)	(1.6 e 8)
	⟨1.04⟩	⟨33.6⟩	⟨0.5⟩	⟨1.1⟩	⟨2.7⟩	⟨0.9⟩	⟨2124⟩
ER^f	3.00	2.47	3.00	8.6	2.88	19.7	−5.42
	[1.8]	[9.0]	[0.6]	[30.3]	[−0.1]	[2.6]	[36.8]
	(0.01)	(0.3)	(0.001)	(0.1)	(0.01)	(0.03)	(2.7)
	⟨0.13⟩	⟨6.3⟩	⟨0.01⟩	⟨1.4⟩	⟨0.2⟩	⟨0.6⟩	⟨42.3⟩
$E(R^e - R^f)$	0.20	0.60	0.02	300.9	0.06	40.1	6.86
	[−25.8]	[93.5]	[−77.6]	[1.2]	[−3.5]	[−3.8]	[3.3 e 7]
	(0.7)	(8.5)	(4.6)	(0.1)	(6.1)	(0.04)	(3.3 e 7)
	⟨20.4⟩	⟨97.9⟩	⟨460.7⟩	⟨1.1⟩	⟨141.3⟩	⟨1.0⟩	⟨2124⟩
σ_q	1.07	2.01	0.31	6.41	0.59	14.2	13.7
	[−5.4]	[26.8]	[−2.6]	[−1.9]	[1.4]	[−0.7]	[7.9]
	(0.1)	(3.0)	(0.1)	(0.1)	(0.8)	(0.01)	(1.1)
	⟨3.5⟩	⟨52.9⟩	⟨1.5⟩	⟨1.4⟩	⟨17.3⟩	⟨0.3⟩	⟨22.0⟩
$\rho(y,q)$	0.40	0.31	0.38	0.33	0.09	0.99	0.47
	[−8.2]	[7.3]	[−7.6]	[1.5]	[3.2]	[−0.2]	[−0.3]
	(0.1)	(1.5)	(0.04)	(0.02)	(0.35)	(0.002)	(0.7)
	⟨2.4⟩	⟨31.2⟩	⟨1.0⟩	⟨0.4⟩	⟨7.6⟩	⟨0.1⟩	⟨14.5⟩
Freq($q < 1$)	24.6	9.2	19.9	20.3	3.8	49.7	31.0
	[−12.5]	[25.5]	[−12.6]	[6.7]	[−2.0]	[−3.8]	[−7.2]
	(0.1)	(3.0)	(0.1)	(0.04)	(0.4)	(0.004)	(0.6)
	⟨2.9⟩	⟨52.6⟩	⟨1.5⟩	⟨0.8⟩	⟨12.6⟩	⟨0.1⟩	⟨13.3⟩
			Panel C - Computation times in seconds				
Time	26.5	28.8	25.6	38.7	19.3	26.5	406

Notes: See next page.

Notes to table 3.

i. Unbracketed numbers: statistic, s^{dp}, based on a single simulation of length 100,000 generated using dynamic programming solution.

ii. Square bracketed numbers: $100 \cdot (\overline{s} - s^{dp})/s^{dp}$, where $\overline{s}$ is the mean of the statistic across $I = 500$ simulated data sets of length $100,000$ observations each. Each of the I datasets was generated by a different conventional PEA solution. For model (7), 48 of the artificial datasets had to be discarded because the capital stock converged to zero.

iii. Round bracketed numbers: Monte Carlo standard error, $100 \cdot \sigma_s/(I \cdot s^{dp})$, for the object in square brackets. Here σ_s is the standard deviation of the statistic across I conventional PEA-generated datasets.

iv. Angular bracketed numbers: coefficient of variation, $100 \cdot \sigma_s/\overline{s}$.

Table 4. Overcoming bias and Monte Carlo variation in conventional PEA[ii]

	conventional PEA		modified conventional PEA[i]		PEA collocation	
	$N=3$	$N=5$	$N=5$	$N=5$	$N=3$	$N=5$
Statistic	$M=10,000$	$M=50,000$	$M=50,000$	$M=10,000$	$M=3$	$M=5$
			Panel A – Quantities			
σ_y	[−0.2]	[0.2]	[0.3]	[0.3]	[0.4]	[0.3]
	(0.04)	(0.03)	(0.02)	(0.07)		
	⟨0.9⟩	⟨0.2⟩	⟨0.003⟩	⟨0.5⟩		
σ_c	[6.3]	[1.0]	[−0.5]	[−0.4]	[−1.1]	[−0.3]
	(0.5)	(0.4)	(0.3)	(0.8)		
	⟨10.6⟩	⟨2.7⟩	⟨2.1⟩	⟨5.9⟩		
σ_i	[−3.8]	[−0.4]	[0.4]	[0.4]	[0.7]	[0.2]
	(0.3)	(0.2)	(0.2)	(0.6)		
	⟨7.3⟩	⟨1.5⟩	⟨1.5⟩	⟨4.3⟩		
$\rho(y,c)$	[−0.2]	[−0.6]	[−0.6]	[−0.8]	[−0.6]	[−0.5]
	(0.1)	(0.1)	(0.1)	(0.2)		
	⟨1.1⟩	⟨0.4⟩	⟨0.5⟩	⟨1.5⟩		
$\rho(y,i)$	[−0.2]	[0.1]	[0.2]	[0.2]	[0.2]	[0.2]
	(0.03)	(0.03)	(0.01)	(0.03)		
	⟨0.7⟩	⟨0.2⟩	⟨0.06⟩	⟨0.2⟩		
			Panel B – Asset Prices and Returns			
$\mathbf{E}R^e$	[1.6 e 8]	[4.7 e 7]	[−13.0]	[−7.5]	[−29.8]	[−8.6]
	(1.6 e 8)	(4.6 e 7)	(8.1)	(24.2)		
	⟨2124⟩	⟨687⟩	⟨65.8⟩	⟨159⟩		
$\mathbf{E}R^f$	[36.8]	[4.2]	[−1.7]	[−0.8]	[−4.2]	[−0.4]
	(2.7)	(1.7)	(−1.6)	(5.0)		
	⟨42.3⟩	⟨11.5⟩	⟨11.8⟩	⟨35.4⟩		
$\mathbf{E}(R^e - R^f)$	[3.3 e 7]	[9.9 e 6]	[−4.1]	[1.0]	[−9.6]	[−2.1]
	(3.3 e 7)	(9.6 e 6)	(2.9)	(8.7)		
	⟨2124⟩	⟨687⟩	⟨21.1⟩	⟨61.1⟩		
σ_q	[7.9]	[−1.9]	[−2.7]	[−5.0]	[−4.9]	[−1.2]
	(1.1)	(0.9)	(1.2)	(3.3)		
	⟨22.0⟩	⟨6.6⟩	⟨8.8⟩	⟨24.6⟩		
$\rho(y,q)$	[−0.3]	[−1.0]	[−1.9]	[−3.5]	[−0.8]	[−1.2]
	(0.7)	(0.9)	(1.0)	(1.8)		
	⟨14.5⟩	⟨6.4⟩	⟨6.9⟩	⟨13.3⟩		
Freq$(q<1)$	[−7.2]	[−4.6]	[−2.0]	[−3.9]	[0.1]	[−1.1]
	(0.6)	(0.9)	(1.0)	(1.8)		
	⟨13.3⟩	⟨6.8⟩	⟨6.9⟩	⟨12.9⟩		
			Panel C - Computation times in seconds			
Time	406	5870	1237	170	0.28	0.66

(i) A version of conventional PEA in which data simulation step (step #1) has been altered to produce greater dispersion. This was done by first computing five values for the capital stock, $k_1, ..., k_5$, based on the zeros of a fifth order Chebyshev polynomial. For each (k_i, θ), we drew 5000 times from $p(\theta' \mid \theta)$, for $i = 1, ..., 5$ and $\theta = -\sigma, \sigma$, respectively. This results in 50,000 sets, (k, θ, θ'), which were used (along with a value for a) to construct $m_2, ..., m_{50,001}$ in the manner described in step #1. These data were then used in the nonlinear regression specified in step #2.

(ii) The entries in the first column are reproduced from the last column in Table 3. There, $I = 500$, though 48 of these had to be discarded because capital converges to zero in simulation. For columns $2-4$, $I = 50$. We did not have to discard any solutions for these cases due to difficulties at the post-solution simulation stage.

Table 5. Computation times for continuous technology implementations of the PEAs

Algorithm	Benchmark Model $\rho = 0.0$	Benchmark Model $\rho = 0.95$
PEA Galerkin	0.22	0.99
(N, M, H)	$(6, 36, 4)$	$(15, 25, 4)$
conventional PEA	5.9	136.9
(N, M)	$(6, 1000)$	$(6, 10000)$

Notes to table 5.

1. Results correspond to a version of the benchmark model in which θ has a Normal distribution. The unbracketed entries are minimum computation times needed to achieve a given level of accuracy, as discussed in the next note. The bracketed numbers correspond to values for the indicated approximation parameters.

2. Minimum computation time: for Chebyshev and conventional PEA we solved the model for various values of N and M. For Chebyshev PEA we selected the values of N and M on this grid which required the smallest computation time, subject to the accuracy constraint that the 11 statistics studied in Tables 3 and 4 are within 10% of the corresponding exact values. For conventional PEA, we selected the values of N and M that minimize computation time, subject to two constraints: bias in each of the 11 statistics studied in Tables 3 and 4 is less than 10% and the coefficient of variation on each statistic is also less than 10% (for these calculations, $I = 30$). Exact solution: approximated by increasing N and M until the 11 statistics implied by each solution procedure are within 1% of each other, and the coefficient of variation implied by conventional PEA is less than 1%.

3. The $N = 6$ implementations of the PEAs included a constant, linear and quadratic terms for each of k and θ and a linear cross term. For the $N = 15$ implementation of Chebyshev PEA we used $\sum_{i=1}^{15} a_i C_i(k, \theta)$, where $C_i(k, \theta)$, $i = 1, ..., 15$ are the elements of the set $\left\{ T_{i1}(\varphi(k)) T_{i2}(\psi(\theta)) \mid \sum_{j=1}^{2} i_j \leq 4 \right\}$. The linear function, ψ, maps $\left[-\overline{\theta}, \overline{\theta}\right]$ into the interval $[-1, 1]$, where $\overline{\theta} = 3\sigma$ and σ is the standard deviation of θ.

Table 6. Computation times and approximation parameters for various algorithms[(i)]

	Model Parameterizations						
Algorithm	(1)	(2)	(3)	(4)	(5)	(6)	(7)
PEA collocation	0.22	0.22	0.27	0.55	0.44	0.27	0.66
(N)	(3)	(3)	(3)	(5)	(5)	(3)	(5)
conventional PEA[(iii)]	26.5	28.8	25.6	38.7	19.3	26.5	406
$(N, M/1000)$	$(3, 10)$	$(3, 10)$	$(3, 10)$	$(3, 10)$	$(3, 10)$	$(3, 10)$	$(3, 10)$
Spectral-Galerkin	1.81	algorithm[(ii)]	1.27	1.32	4.23	4.01	algorithm[(ii)]
$(N_\sigma, N_{-\sigma}, M)$	$(8, 4, 30)$	failed	$(8, 4, 30)$	$(8, 4, 30)$	$(10, 5, 100)$	$(10, 5, 50)$	failed
FEM collocation[(iv)]	63.6	262	55.1	31.7	231	207	876
(N)	(24)	(108)	(72)	(72)	(72)	(36)	(124)
FEM Galerkin[(iv)]	63.7	2430	215	1536	58.0	22.2	144
(N)	(36)	(36)	(36)	(36)	(36)	(6)	(36)

Notes to table 6.

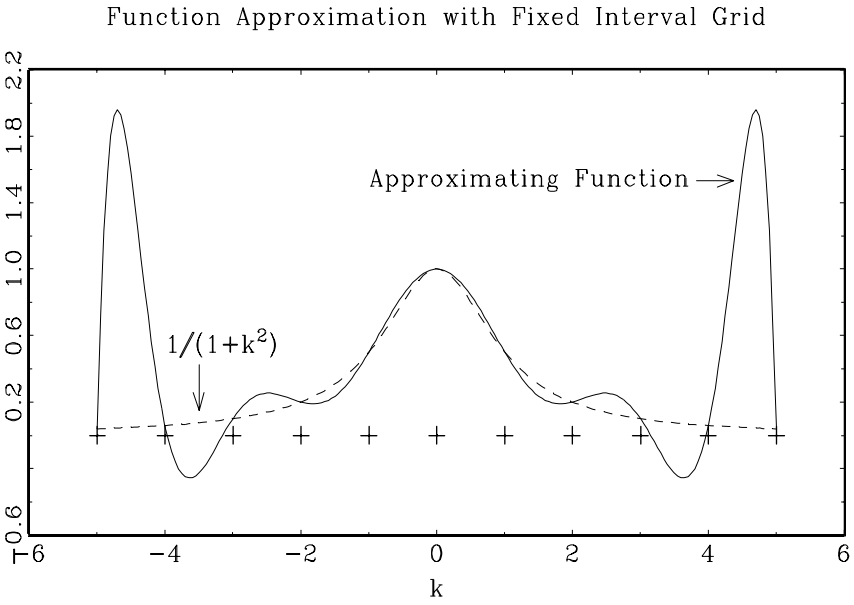
(i) With the exception of results for conventional PEA, numbers not in parentheses are minimal computation times, in seconds, needed by various computational strategies to achieve a given degree of accuracy. Minimal computation times and algorithm parameters were chosen in the same way as described in Table 5, note 2. In particular, the required degree of accuracy is that a model solution must imply a set of values for the 11 statistics studied in Tables 3 and 4 that come within at least 10% of the corresponding exact values obtained by dynamic programming.

(ii) The quasi-Newton algorithm crashed for all algorithm parameters considered.

(iii) The conventional PEA entries correspond to the fastest of the $I = 500$ solutions generated for each model parameterization.

(iv) The parameter N refers to the number of parameters in the policy function. For further details, see the appendix.

Figure 1. Alternative methods of function approximation



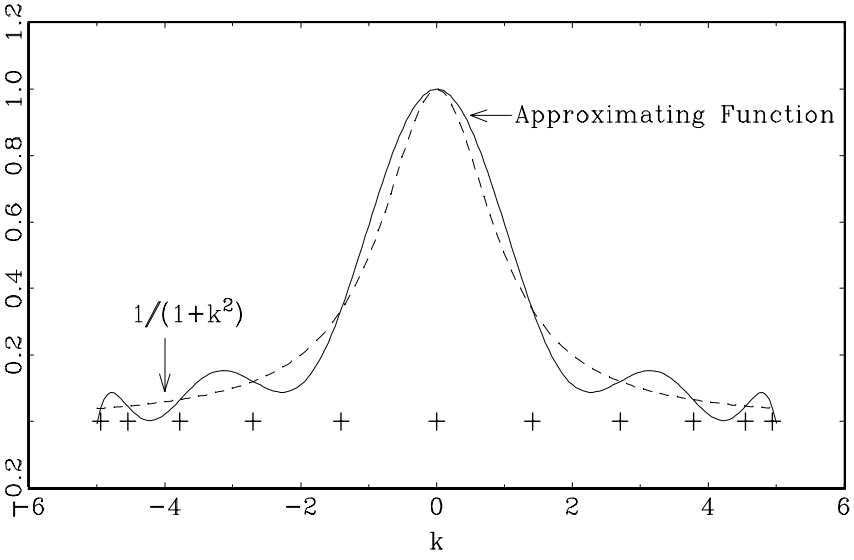
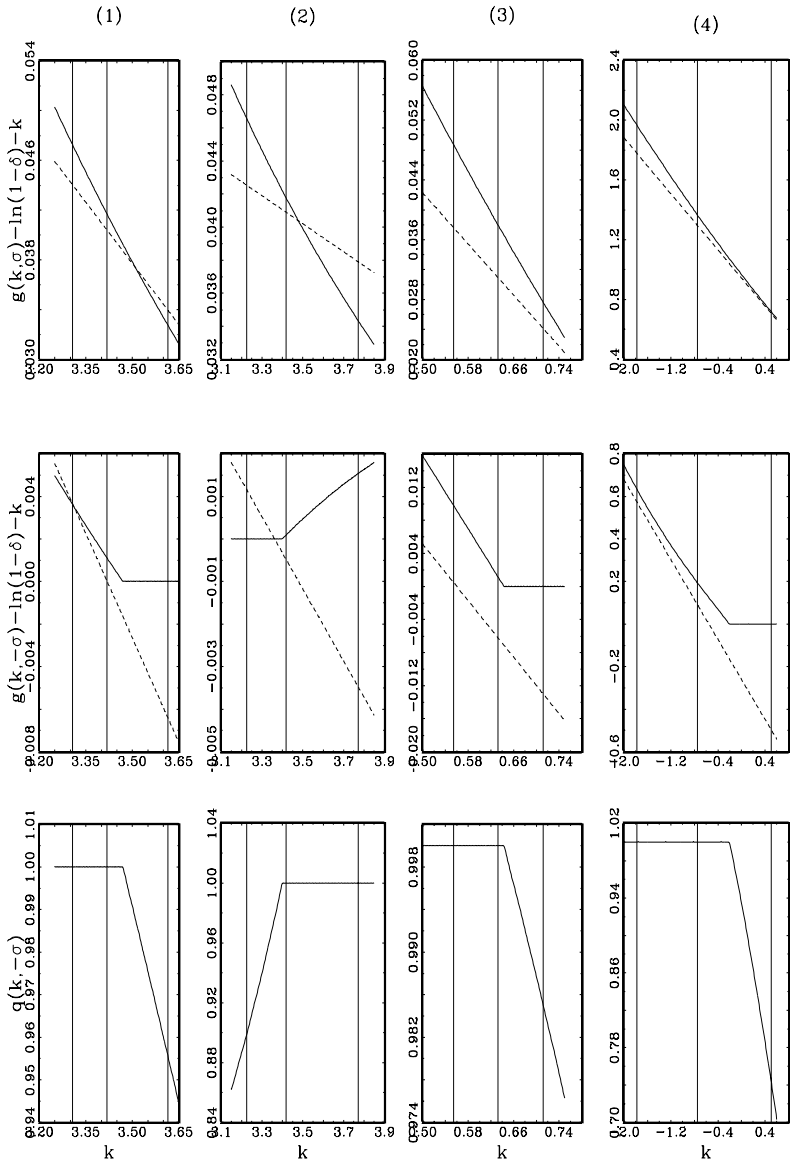


Figure 2. Policy functions for models (1)-(4)

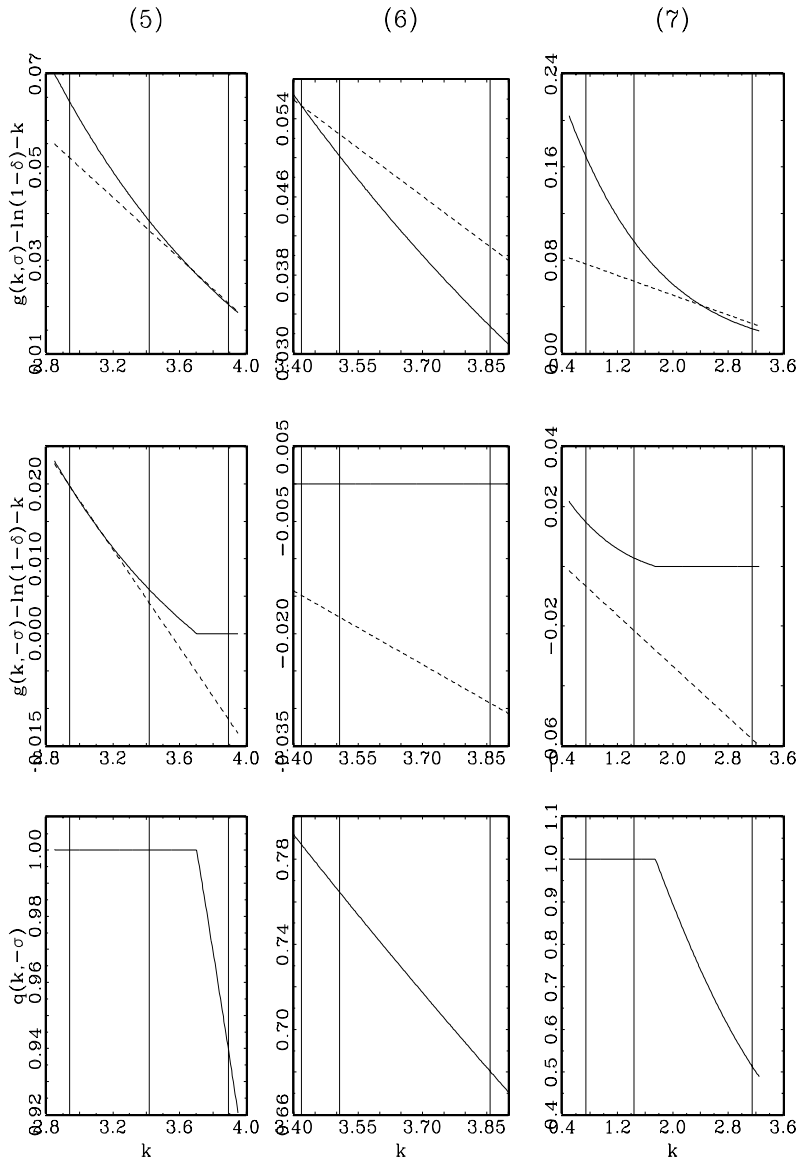


Notes:

i. In each plot the middle vertical line indicates the non-stochastic steady state value of k; the other two vertical lines define a symmetric 95% confidence interval for k.

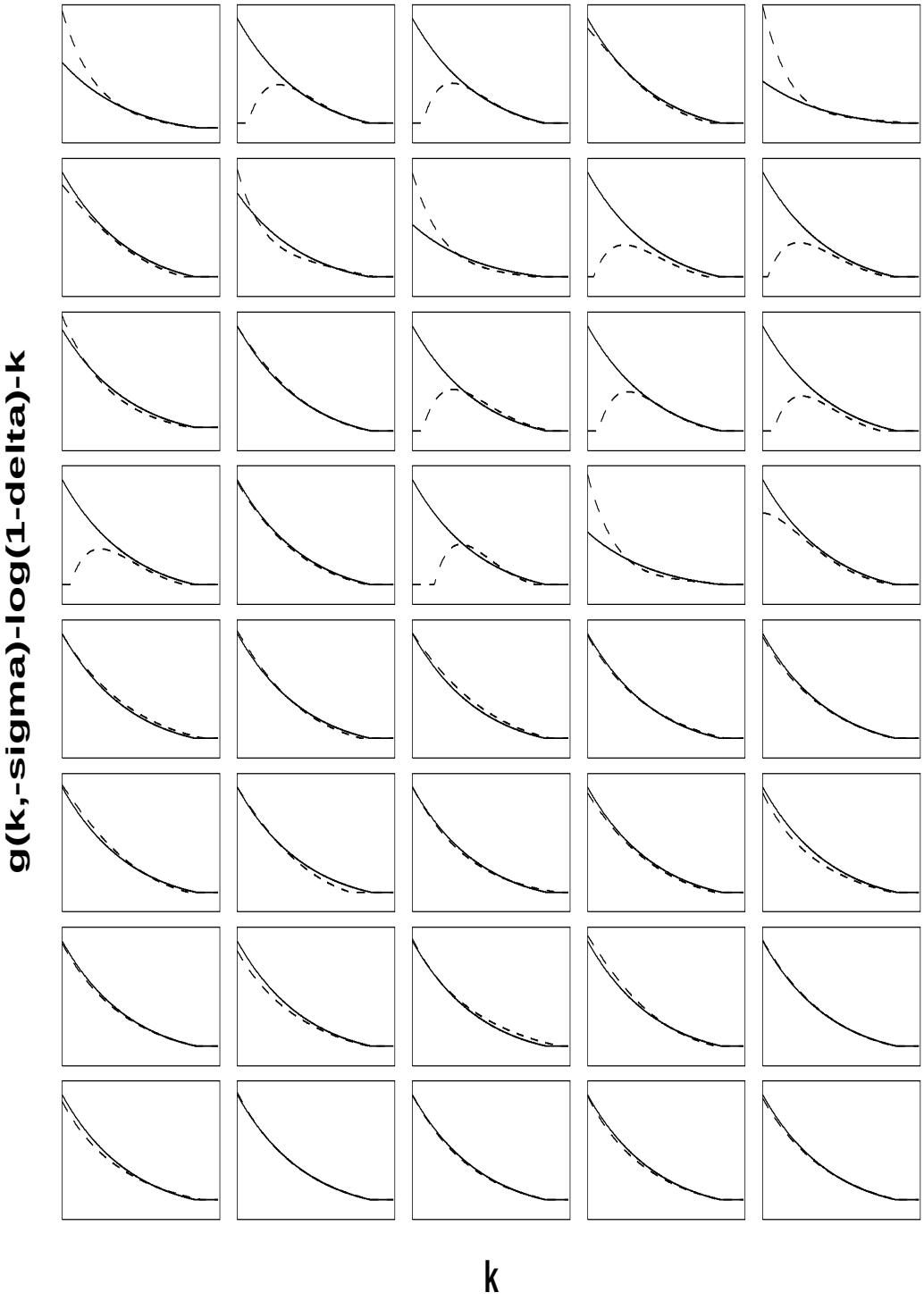
ii. The solid lines in each plot are exact solutions for the indicated policy function; in the top two rows the dashed line is a log-linear approximation to the indicated policy function.

Figure 3. Policy functions for models (5)-(7)



Notes: See the notes to figure 2.

Figure 4. Conventional PEA and Modified Conventional PEA approximations of $g(k, -\sigma) - \log(1 - \delta) - k$, model(7)



Note: In each plot the solid line is the exact solution and the dashed line is one solution computed using a PEA. The first four rows plot Conventional PEA solutions and the last four rows plot Modified Conventional PEA solutions.

Figure 5. Endogenous and exogenous capital stock distributions

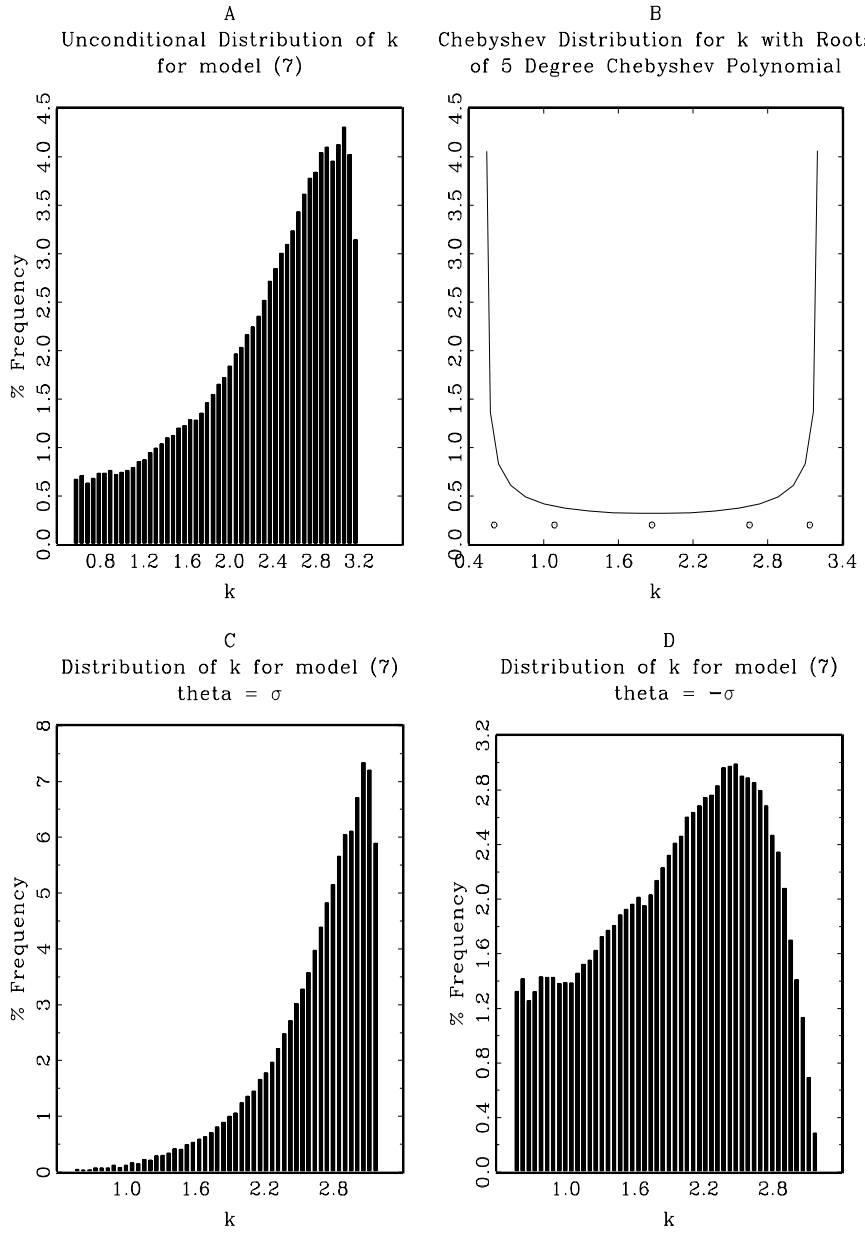


Figure 6. Euler errors for PEA collocation on model (7)

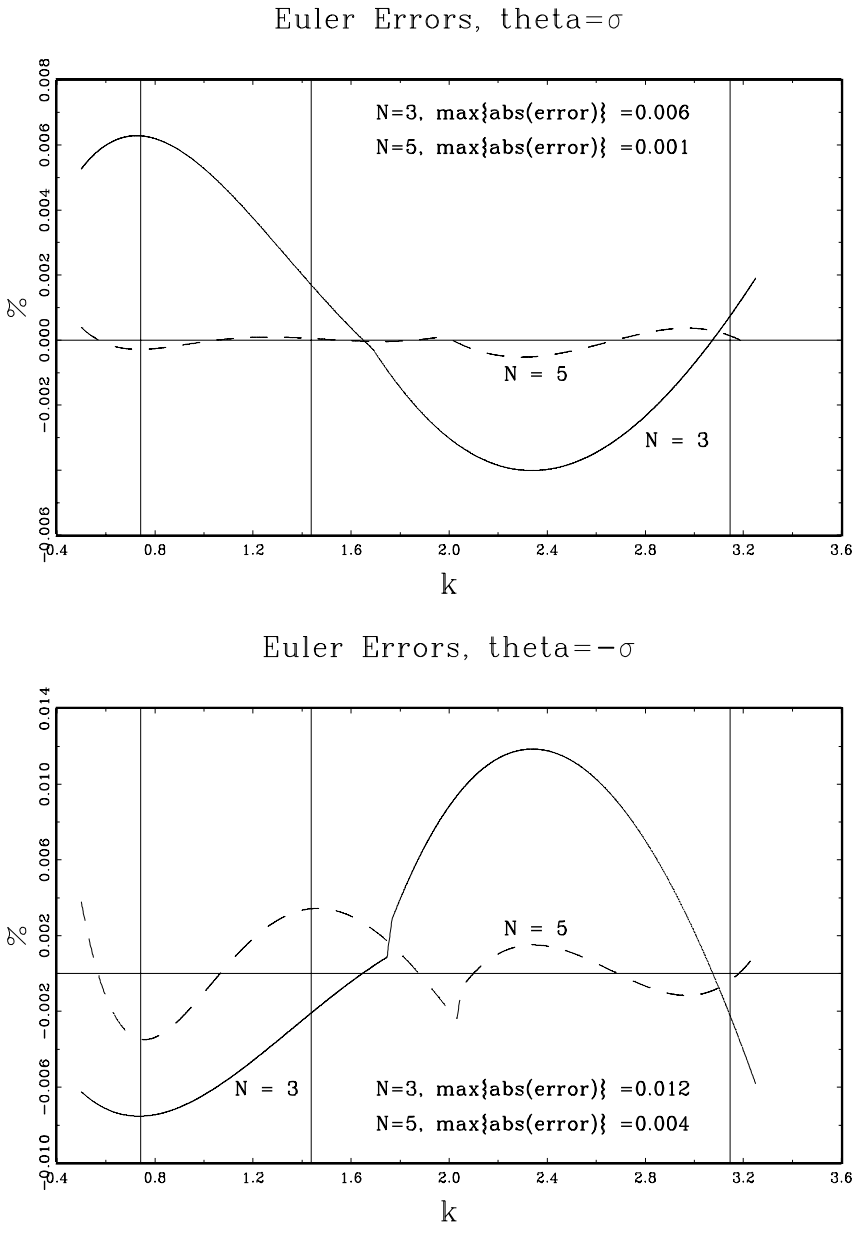


Figure 7. PEA collocation with the Wright-Williams conditional expectation