Working Paper Series

Algorithms for Solving Dynamic Models with Occasionally Binding Constraints

Lawrence J. Christiano and Jonas D.M. Fisher

Working Papers Series
Macroeconomic Issues
Research Department
Federal Reserve Bank of Chicago
April 1994 (WP-94-6)

FEDERAL RESERVE BANK
OF CHICAGO

Algorithms for Solving Dynamic Models with Occasionally Binding Constraints*

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

April, 1994

Abstract

We describe several methods for approximating the solution to a model in which inequality constraints occasionally bind, and we compare their performance. We apply the methods to a particular model economy which satisfies two criteria: it is similar to the type of model used in actual research applications, and it is sufficiently simple that we can compute what we presume is virtually the exact solution. We have two results. First, all the algorithms are reasonably accurate. Second, on the basis of speed, accuracy and convenience of implementation, one algorithm dominates the rest. We show how to implement this algorithm in a general multidimensional setting, and discuss the likelihood that the results based on our example economy generalize.

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

Send correspondence to: Prof. Jonas D.M. Fisher, Department of Economics, Social Science Centre, University of Western Ontario, London, ON, Canada, N6A 5C2.

*We thank Graham Candler, Kenneth Judd, Albert Marcet, David Marshall, and Ellen McGrattan for helpful comments. Christiano is grateful to the National Science Foundation for support. 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 Banks of Chicago and Minneapolis.
‡University of Western Ontario.

1. Introduction

In recent years there has been substantial 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. The main 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] An important consideration in selecting algorithms for solving models like these is the quantity of computer and programmer time required to achieve an acceptable degree of accuracy. The purpose of this paper is to help shed light on these issues.

We describe six algorithms, and evaluate their accuracy in solving the one-sector infinite horizon optimal growth model with random productivity disturbances. 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 four 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 with the Euler equation associated with the recursive

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

[2] For an example of the former, see Atkeson and Kehoe (1993), and Boldrin, Christiano and Fisher (1994). Examples of the latter include 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.

representation of the model. Thus, a solution is viewed as a policy function relating decisions to a small number of state variables. All but one of the algorithms considered work with a version of the model in which the nonnegativity constraint is incorporated by the method of Lagrange multipliers. These include suitably modified versions of the algorithms emphasized by Bizer and Judd (1989), Coleman (1988), Danthine and Donaldson (1981), Judd (1992a), and Marcet (1988). The sixth algorithm, an example of the Finite Element Method, works with a version of the model in which the nonnegativity constraint is incorporated by a penalty function method. This algorithm has been advocated by McGrattan (1993).[4]

Our main finding is that, for the model economy studied, one algorithm dominates the others in terms of speed, accuracy and programmer time. This algorithm approximates the solution indirectly by parameterizing the conditional expectation in the Euler equation using an exponentiated polynomial, as in Marcet (1988).[5] We show that conventional implementations of parameterized expectations have some shortcomings, and document that our preferred algorithm dominates on these dimensions.

In our example, there are two principal advantages in parameterizing a conditional expectation. First, the conditional expectation function is smoother than other functions characterizing the solution, such as the policy function. In general, it is easier to obtain an accurate approximation, the smoother the function being approximated. A second advantage is that working with parameterized expectations is efficient from the point of view of programmer time. In the context of methods based on Lagrange multipliers, the requirement that the Euler equations and Kuhn-Tucker conditions be satisfied implies a convenient mapping from a parameterized expectation function into policy and multiplier functions. This obviates the need to separately parameterize the latter. Methods which focus on the policy function must jointly parameterize the policy and multiplier functions. Doing this

[4]See the chapter in Judd (1992b) on rational expectations models for references to earlier analyses of models with nonnegativity constraints. As the material in that chapter indicates, several of the methods used in this paper actually correspond to approaches taken by Gustafson and other agricultural economists decades ago.

[5]For other applications of the PEA when there are occasionally binding constraints, see den Haan (1993), Marcet and Ketterer (1989), Marcet and Marimon (1992), Marcet and Singleton (1990), and McCurdy and Ricketts (1992).

in a way that the Kuhn-Tucker conditions are satisfied is tricky and adds to programmer time. For methods that focus on policy functions, an alternative to working with Lagrange multipliers is to work with a penalty function. However, these methods require searching for a parameter in the penalty function, which can add substantially to programmer and computer time. Although we carefully document these statements for our model economy, we expect them to be true in a broader class of models as well.

The paper is organized as follows. In the following section the model to be solved is described. This is followed by a review of how the six algorithms can be used to approximate the unconstrained version of the model in which the nonnegativity constraint on investment is ignored. In the next section we describe the way these algorithms are modified to accommodate the nonnegativity constraint on investment. Results from implementing the algorithms for a particular parameterization of the model are discussed in section 5. In the final section we offer some concluding remarks.

2. The Model

We examine a simple version of the stochastic growth model with inelastic labor supply. 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 and $\beta \in (0,1)$ is the agent's discount factor. The aggregate resource constraint is given by

$$c_t + k_{t+1} - (1-\delta)k_t \leq f(k_t, \theta_t) \equiv \exp(\theta_t) k_t^{\alpha}, \tag{2.1}$$

where k_t denotes the beginning-of-period-t stock of capital, and δ, $\alpha \in (0,1)$. Here, δ is the rate of depreciation of capital, and α is capital's share in production. We assume $\theta_t \in \Theta$ is exogenous with respect to k_t and 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 our computational experiments, we assume θ_t is *i.i.d.* with $\Theta \equiv \{\sigma, -\sigma\}$, and that the probability associated with each of σ and $-\sigma$ is 1/2. The initial stock of capital, k_0, is given.

In the irreversible investment version of the model, we require that gross investment be non-negative, *i.e.*:

$$k_{t+1} - (1-\delta)k_t \geq 0. \tag{2.2}$$

In the reversible investment version, (2.2) is ignored.

We formulate the planner's problem in recursive form. In doing so we drop t subscripts and use $'$ to denote next period's value of a variable. The planner's dynamic program is then given by

$$W(k,\theta) = \max_{f(k,\theta)+(1-\delta)k \geq k' \geq 0} U(c(k,k',\theta)) \; + \beta \int_{\theta' \in \Theta} W(k',\theta') p_{\theta'}(\theta' \mid \theta) d\theta'. \tag{2.3}$$

Equation (2.2) must also be satisfied in the irreversible investment version of the model. Assumptions we will place on $U(\cdot)$ guarantee that (2.1) always binds for this economy. In (2.3) we have used this fact to replace consumption in $U(\cdot)$ with $c(\cdot)$, the function implicit from (2.1). Finally, $W(\cdot)$ is the planner's value function.

To solve the planner's problem we introduce a Lagrange multiplier, λ, on constraint (2.2). The solution to the planner's problem is a set of time invariant functions $g(k,\theta)$ and $h(k,\theta)$ that determine k' and λ, respectively, given values of k and θ. These functions must satisfy an Euler equation,

$$U_c(k, g(k,\theta), \theta) - h(k,\theta) - \beta \int_{\theta'} m(k,\theta,\theta'; g, h) p_{\theta'}(\theta' \mid \theta) d\theta' = 0, \tag{2.4}$$

and a set of Kuhn-Tucker conditions

$$g(k,\theta) - (1-\delta)k \geq 0, h(k,\theta) \geq 0, \text{and } \; h(k,\theta)[g(k,\theta) - (1-\delta)k] = 0. \tag{2.5}$$

Also,

$$\begin{aligned} m(k,\theta,\theta'; g, h) = U_c\left(g(k,\theta), g\left(g(k,\theta),\theta'\right),\theta'\right)\left[f_k\left(g(k,\theta),\theta'\right) + 1 - \delta\right] \\ -h\left(g(k,\theta),\theta'\right)(1-\delta). \end{aligned} \tag{2.6}$$

In (2.4)-(2.6), f_k denotes the derivative of f, while U_c combines the derivative of U with the

function $c(k, k', \theta)$.

If the standard deviation of the technology shock, σ, is small enough, (2.2) will never bind and $\lambda = 0$ for all θ and k in the ergodic set for capital. Methods for approximating the solution to the planner's problem are well known for this case. Here we consider the case where σ is large enough so that $\lambda > 0$ with nonzero probability.

3. Solving the Unconstrained Model

It is convenient for us to begin by reviewing how the six algorithms studied in this paper are implemented in the reversible investment version of the model. To have a consistent terminology for discussing and comparing the algorithms, we use the framework in Reddy (1993)'s numerical analysis text, which corresponds closely to the framework presented in Judd (1992a,1992b). With one class of exceptions, the algorithms considered in this paper are what Reddy calls *weighted-residual methods*. The exceptions, standard implementations of Marcet (1988)'s *Parameterized Expectations Algorithm* (PEA), fail to be weighted residual methods only because of a technicality.

In the reversible investment case, there is only (2.4) with $h \equiv 0$ to solve, or, more compactly,

$$R(k, \theta; g) \equiv \frac{U_c(k, g(k, \theta), \theta)}{\beta} - \int_{\theta'} m(k, \theta, \theta'; g, 0) p_{\theta'}(\theta' \mid \theta) d\theta' = 0, \tag{3.1}$$

for $k \geq 0$ and all $\theta \in \Theta$. In (3.1), the 0 argument in m reflects that in the reversible investment case, the multiplier on gross investment, h, is identically zero. We refer to $R(k, \theta; g)$ as the *Euler residual* function. Solving the model amounts to finding a function g that solves the functional equation, $R(g) = 0$, *i.e.*, sets the Euler residual function to zero everywhere.[6] This problem is complicated by the fact that k can take on a continuum of values. This implies that solving (3.1) is a problem of solving a continuum of equations (one for each k, θ)

[6] Uniqueness of this solution requires also imposing a transversality condition. The approximate solutions computed for this paper all trivially satisfy the transversality condition, because they imply a bounded ergodic set for capital.

in a continuum of unknowns (one k' value for each k, θ). Apart from a few cases, in which R has a convenient structure, solving this problem is computationally infeasible.

Instead, we select a function, $\hat{g}_a$, parameterized by a finite set of coefficients, a, and choose values for a, a^*, to make $R(\hat{g}_a)$ 'small'. Weighted-residual methods compute a^* as the solution to what Reddy (1993, p. 580) refers to as the *weighted-residual form* of (3.1):[7]

$$\int_{\theta \in \Theta} \int_{k \geq 0} R(k, \theta; \hat{g}_{a^*}) w(k, \theta) dk d\theta = 0. \tag{3.2}$$

The choice of weighting functions in (3.2), $w(k, \theta)$, operationalizes the notion of 'small'. In practice, the range of integration over k in (3.2) is finite, with $\underline{k} \leq k \leq \overline{k}$. The boundary points of this interval are chosen to ensure that the values of the capital stock generated by the model always lie in the interior of the interval, $(\underline{k}, \overline{k})$. Computationally, we obtain an initial guess of $\underline{k}$ and $\overline{k}$ by finding the interval $(\underline{k}, \overline{k})$ that contains k_0 and the ergodic set of the log-linear approximation to the policy function in its interior.[8]

To apply the weighted-residual method, one has to select a family of approximating functions, $\hat{g}_a$, a set of weighting functions, $w(k, \theta)$, and strategies for evaluating the integrals in (3.2) and the integral implicit in the expectation operator in R. The procedures we consider make different choices on these three dimensions. Two general types of $\hat{g}_a$ functions include *spectral* and *finite element* functions. In the former, each component of a influences $\hat{g}_a(k, \sigma)$, or $\hat{g}_a(k, -\sigma)$, over the whole range of k while in the latter, each component of a has influence over only a limited range of k's.[9] Regarding the weighting functions, a necessary condition for (3.2) to pin down the parameter vector a, is that there be a number of weighting functions equal to the dimension of a. We consider three types of weighting functions. In one, the $w(k, \theta)$'s are related to the basis functions generating $\hat{g}_a(k, \theta)$, in which case the algorithm is an example of the *Galerkin* method. In another, the basis functions are

[7] In our case, (3.2) reduces to $\int_{k \geq 0} R(k, \sigma; \hat{g}_a) w(k, \sigma) dk + \int_{k \geq 0} R(k, -\sigma; \hat{g}_a) w(k, -\sigma) dk = 0$.

[8] See Christiano (1991, Appendix) for details about solving the model studied here using a log-linearization method.

[9] See Judd (1992a,1992b), McGrattan (1993), and Reddy (1993) for more detailed discussions of spectral and finite-element functions, respectively.

particular kinds of dirac delta functions, in which case the algorithm is an example of the *collocation* method. The versions of the PEA that we consider also choose an a to solve an expression of the form (3.2).[10] However, technically, standard implementations of the PEA do not fall in the class of weighted-residual methods as defined by Reddy. This is because, as we will see below, they work with weighting functions in which the parameter vector, a, is an argument. Finally, two numerical procedures are used to evaluate the integrals in (3.2): quadrature methods and Monte Carlo integration. We now turn to a detailed discussion of the six algorithms considered.

3.1. Two Spectral, Weighted-Residual Methods

3.1.1. Parameterized Expectations

What distinguishes the class of Parameterized Expectations Algorithms is that they approximate the function, g, indirectly by approximating the conditional expectation in (3.1) as follows:

$$\int_{\theta'} m(k,\theta,\theta';g,0)p_{\theta'}(\theta' \mid \theta)d\theta' \approx \exp[\widehat{\epsilon}_a(k,\theta)], \tag{3.3}$$

where $\widehat{\epsilon}_a(k,\theta)$ is a function, parameterized by a finite set of parameters, a. The purpose of the exponential in (3.3) is to guarantee nonnegativity. The PEA's approximation of g is obtained by solving $U_c(k,k',\theta) = \beta \exp(\widehat{\epsilon}_a(k,\theta))$ for k' given each k,θ, yielding:

$$\widehat{g}_a(k,\theta) = \exp(\theta)k^\alpha + (1-\delta)k - U_c^{-1}\left[\beta \exp(\widehat{\epsilon}_a(k,\theta))\right], \tag{3.4}$$

where $U_c^{-1}[\cdot]$ denotes the inverse function of U_c. The PEA approximation of h, $\widehat{h}_a$, is trivial when the nonnegativity constraint on investment is ignored. We simply set $\widehat{h}_a(k,\theta) = 0$ for all k,θ. It remains to describe how the various PEA's go about computing a^*.

One way to view the three versions of the PEA that we consider, is to think of them as solving a particular fixed point problem. A given value of a induces, via (2.6), (3.4), and $p_{\theta'}$,

[10] For a related discussion, see Marcet and Marshall (1994).

a distribution on $m(k,\theta,\theta';\widehat{g}_a,\widehat{h}_a)$ for any fixed (k,θ). A new set of parameters values, a', is found which makes $R(k,\theta;\widehat{g}_{a'}) = \exp(\widehat{e}_{a'}(k,\theta)) - \int_{\theta'} m(k,\theta,\theta';\widehat{g}_a,\widehat{h}_a)p_{\theta'}(\theta' \mid \theta)d\theta'$ close to zero in the sense of (3.2).[11] Denote this mapping from a to a' by $a' = S(a;N^p)$. The value of a selected by the PEA is the fixed point, a^*, such that $a^* = S(a^*;N^p)$. We computed a^* by applying a standard nonlinear equation solving routine.[12] The versions of the PEA that we consider differ in the form of the weighting functions used in (3.2) and in the computational strategy for evaluating the integrals.

Conventional PEA

In our implementation of what we call *conventional PEA*, we parameterize the expectation function as follows:

$$\widehat{e}_a(k,\theta) \equiv \sum_{i=0}^{N^p-1} a_i(\theta)P_i(\varphi(k)), \tag{3.5}$$

for $\theta \in \Theta$. The basis functions for $\widehat{e}_a$ are the N^p Legendre polynomials, $P_i(\cdot)$, $i = 0,\ldots,N^p-1$.[13] Here, a is the $2N^p \times 1$ dimensional vector of parameters, $\{a_i(\theta), i = 0,\ldots,N^p-1, \theta \in \Theta\}$. We define $\varphi : (\underline{k},\overline{k}) \to (-1,1)$ to ensure the polynomials in (3.5) are of similar orders of magnitude. That is,

$$\varphi(k) = 2\frac{k-\underline{k}}{\overline{k}-\underline{k}} - 1. \tag{3.6}$$

[11]Note, this definition of R coincides exactly with the one in (3.1), since the PEA policy function, $\widehat{g}_a$, implies $U_c(k,\widehat{g}_a(k,\theta),\theta) \equiv \beta \exp(\widehat{e}_a(k,\theta))$.

[12]The standard implementation of the PEA finds a^* by a method of successive approximation, as the limit of $a, S(a;N^p), S^2(a;N^p),\ldots$. As noted by Judd (1992b, chapter 13, pp. 11-14) and Marcet (1988), this algorithm can yield explosive, oscillatory sequences, $a, a',\ldots$, particularly for higher values of N^p. One alternative is to instead iterate on the operator $\widetilde{S}$, where $\widetilde{S}(a) = (1-\mu)a + \mu S(a;N^p)$, 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, the equation-solving alternative described in the text finds a^* more quickly and reliably. The nonlinear equation solver we used is NLSYS in GAUSS. See Marcet (1988), Marshall (1992), and 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 $N^p \to \infty$.

[13]The polynomials, P_i, have domain and range $(-1,1)$, and are defined as follows. 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 $i \geq 1$. The orthogonality property of these polynomials is designed to mitigate multicollinearity problems associated with step 2 of the conventional PEA and PEA with exogenous oversampling, which is discussed later in the text. This construction of our basis polynomials may mitigate multicollinearity, but does not eliminate it, since that requires that the integrand in the above orthogonality condition be weighted by a probability density for x.

For any given value of a, a' is found by running a nonlinear regression of $m(k,\theta,\theta';\widehat{g}_a,\widehat{h}_a)$ on the space of functions generated by $\exp(\widehat{e}_{\widetilde{a}}(k,\theta))$for $\widetilde{a}\in\Re^{2N^p}$. To specify the regression, we need to indicate how many observations of every possible type, (k,θ,θ'), were used. This is accomplished by specifying a density function, $p(k,\theta,\theta';a)$. This density has the following structure. Let $p_1(k,\theta;a)$ denote the marginal density of (k,θ), which may depend on a. Then, $p(k,\theta,\theta';a)=p_1(k,\theta;a)p_{\theta'}(\theta'\mid\theta)$. The nonlinear regression is:

$$\begin{aligned} a' &= \underset{\widetilde{a}\in\Re^{2N^p}}{\operatorname{argmin}} \int_{k,\theta,\theta'} \left\{\exp(\widehat{e}_{\widetilde{a}}(k,\theta)) - m(k,\theta,\theta';\widehat{g}_a,\widehat{h}_a)\right\}^2 p(k,\theta,\theta';a)dkd\theta d\theta' \\ &\equiv S(a;N^p). \end{aligned} \tag{3.7}$$

The first order conditions associated with this regression are:

$$\int_{k,\theta,\theta'} \left[\exp(\widehat{e}_{a'}(k,\theta)) - m(k,\theta,\theta';\widehat{g}_a,\widehat{h}_a)\right]\exp(\widehat{e}_{a'}(k,\theta))\frac{d\widehat{e}_{a'}(k,\theta)}{da_l}p(k,\theta,\theta';a)dkd\theta d\theta' = 0, \tag{3.8}$$

for $l=1,\ldots,2N^p$. Taking into account the structure of p, the fixed point, a^*, is easily seen to solve the version of (3.2) with weighting matrices

$$w^l(k,\theta;a^*) = p_1(k,\theta,a^*)\exp(\widehat{e}_{a^*}(k,\theta))\frac{d\widehat{e}_{a^*}(k,\theta)}{da_l}, \tag{3.9}$$

for $l=1,\ldots,2N^p$.

Under conventional PEA, all three integrals in the function S are evaluated by the following Monte Carlo method. First simulate a series of length T, $\{\theta_0,\theta_1,\ldots,\theta_T\}$, using a random number generator, and compute an initial value of a.[14] Then:

1. Simulate $\{k_1,k_2,\ldots,k_{T+1}\}$ recursively using $k_{t+1}=\widehat{g}_a(k_t,\theta_t)$, $t=0,1,\ldots T$ and the given initial value, k_0.

[14]One way to obtain a starting value for a is to generate the data according to step 1 of the conventional PEA algorithm using the version of $\widehat{g}_a$ based on log-linear approximation, and taking the starting value of a as the one that solves (3.10).

2. Find a', in the nonlinear least-squares regression problem:[15]

$$a' = \underset{\widetilde{a}\in\Re^{2NP}}{\operatorname{argmin}} \frac{1}{T}\sum_{t=0}^{T-1}\left[\exp(\widehat{e}_{\widetilde{a}}(k_t,\theta_t)) - m(k_t,\theta_t,\theta_{t+1};\widehat{g}_a,\widehat{h}_a)\right]^2 \tag{3.10}$$

for $l = 1,\ldots,2N^P$.

For T large, the function being minimized in (3.10) coincides with the one being minimized in (3.7).

The PEA specification of the density of (k,θ) concentrates observations on points of high probability. Our computational experiments suggest that greater dispersion in (k,θ) may be desirable. Marcet and Marimon (1992) have made this observation in the context of a study of the far-from-steady-state properties of a model. However, we find that this may be true even when the objects of interest are properties of the steady state distribution of functions of k. In our example, by increasing dispersion relative to conventional PEA, one gets a more accurate estimate of properties of the steady state distribution of k using a lower value of T. Presumably, this reflects the well-known fact that high variance in explanatory variables implies greater precision in regression estimates.[16] With these considerations in mind, we studied two perturbations of the conventional PEA which imply greater dispersion in (k,θ).

PEA with Exogenous Oversampling

Under Marcet and Marimon's (1992) *PEA with exogenous oversampling*, $p_1(k,\theta;a)$ is modified so that extra mass is exogenously placed in particular regions of the state space,

[15]The nonlinear least squares problem in step 2 was handled using a version of the procedure applied in Marshall (1992). Let $\widehat{S}(a;N^P)$ denote $S(a;N^P)$ with (3.10) in step 2 replaced by

$$a' = \underset{\widetilde{a}\in\Re^{2NP}}{\operatorname{argmin}} \frac{1}{T}\sum_{t=0}^{T-1}\left[\widehat{e}_{\widetilde{a}}(k_t,\theta_t) - \log\left(m(k_t,\theta_t,\theta_{t+1};\widehat{g}_a,\widehat{h}_a)\right)\right]^2 = 0,$$

for $l = 1,\ldots,2N^P$. This is just a linear regression and is easy to solve. We first computed a^{**} such that $a^{**} = \widehat{S}(a^{**};N^P)$ using a non-linear equation solver. Finding a^{**} takes little computer time because of the simplicity of the modified step 2. We then used a^{**} as a starting value for solving $a^* = S(a^*;N^P)$. In our experience, a^* and a^{**} are very close.

[16]In our context, there is an offsetting effect. Namely, with too much dispersion, accuracy of the parameterized expectation in the neighborhood of steady state is sacrificed.

(k, θ). This method is implemented by adding J terms to the criterion function in step 2 of the conventional PEA. These terms require simulating J sequences of length $\tilde{T}$ each, of the technology shock: $\{\theta_t^j, t = 0, ..., \tilde{T} - 1\}$ and of the capital stock, $\{k_t^j, t = 1, ..., \tilde{T}\}$, for $j = 1, ..., J$, where k_0^j is a value of the capital stock close to the region of interest. The additional terms are:

$$\sum_{j=1}^{J} \left[\sum_{t=0}^{\tilde{T}-1} \left\{ \exp\left(\hat{e}_{a'}(k_t^j, \theta_t^j) \right) - m(k_t^j, \theta_t^j, \theta_{t+1}^j; \hat{g}_a, \hat{h}_a) \right\}^2 \right]. \tag{3.11}$$

PEA-Collocation

Once the PEA is expressed as a weighted residual method, it is clear that there are many other ways to find a^*. One could evaluate all integrals using one of a variety of available quadrature formulas.[17] Also, there are a variety of different weighting schemes that one can use, some of which are discussed below. Finally, there are a great many alternative classes of finite parameter functions that one can use to parameterize expectations.

Here, we pursue one particularly promising weighted residual method. It works with a more dispersed set of (k, θ)'s than does conventional PEA. It converts the nonlinear regression in conventional PEA and PEA with exogenous oversampling into a linear regression on an orthogonal set of explanatory variables. There is reason to expect (and this is confirmed in section 5) that the number of observations required in the regression is very small. Finally, there is some *a priori* reason for believing that the method may have good accuracy properties.

The method we pursue is a collocation method, in which the weighting functions are dirac delta functions. It is consistent with the use of either Monte Carlo or quadrature methods to evaluate the integral in the definition of R. The dirac delta functions are constructed to assign positive weight to the values of k corresponding to the N^p zeros of the N^p'th order

[17] Quadrature methods approximate integrals by the weighted sum of the integrand, evaluated at a relatively small number of points. This approximation to integrals is known to be very accurate in the one dimensional case, and recently Judd and Bernardo (1994), applying the ideas of Stroud (1971), have argued that multidimensional quadrature integration can also be made very efficient.

Chebyshev polynomial, T_{N^p}.[18] That is, given a we compute a' to solve the following problem:

$$R(k_i, \theta; \hat{g}_{a'}) = \exp\left(\hat{e}_{a'}(k_i, \theta)\right) - \int_{\theta'} m(k_i, \theta, \theta'; \hat{g}_a, \hat{h}_a) p(\theta' \mid \theta) d\theta' = 0 \tag{3.12}$$

for $i = 1, ..., N^p$, $\theta \in \Theta$. Here, $k_i = \varphi^{-1}(r_i)$, where $T_{N^p}(r_i) = 0$, $i = 1, \ldots, N^p$. In addition, we replaced P_i in (3.5) with Chebyshev polynomials, T_i, $i = 0, 1, \ldots, N^p - 1$. There is a slight abuse of notation in (3.12), since R is also a function of a, but our notation does not reflect this. The system of equations, (3.12), is (3.2) with the weighting functions, $w(k, \theta)$, constructed using delta functions, $\delta(k - k_i)$, $i = 1, \ldots, N^p$ as follows:

$$\overline{w}^i(k, \theta) = \begin{cases} \delta(k - k_i) & \text{for } \theta = \sigma \\ 0 & \text{for } \theta = -\sigma \end{cases}, \tag{3.13}$$

and

$$\underline{w}^i(k, \theta) = \begin{cases} \delta(k - k_i) & \text{for } \theta = -\sigma \\ 0 & \text{for } \theta = \sigma \end{cases}, \tag{3.14}$$

$i = 1, \ldots, N^p$.

Our choice of Chebyshev polynomials as basis functions for the parameterized expectations and for determining the grid on k was influenced by the following two considerations. First, the discrete orthogonality property of Chebyshev polynomials greatly facilitates the computations in (3.12) when N^p is large.[19] This property implies that the mapping, $a' = S(a; N^p)$, defined by the solution to (3.12), has a particularly simple analytic form:

$$a_l(\theta)' = \frac{\mu}{N^p} \sum_{i=1}^{N^p} T_l\left(\varphi(k_i)\right) \log\left[\int_{\theta'} m(k_i, \theta, \theta'; \hat{g}_a, \hat{h}_a) p(\theta' \mid \theta) d\theta'\right], \tag{3.15}$$

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

[19] The discrete orthogonality property is that, for $i, j < N^p$:

$$\sum_{k=1}^{N^p} T_i(r_k) T_j(r_k) = \begin{cases} 0, \text{for } i \neq j \\ N^p, \text{ for } i = j = 0 \\ N^p/2, \text{ for } i = j \neq 0, \end{cases}$$

where r_k, $k = 1, \ldots, N^p$ are the roots of $T_{N^p}(\cdot)$.

for $l = 0, 1, \ldots, N^p - 1$, $\theta \in \Theta$. Here, $\mu = 2$ for $l > 0$ and $\mu = 1$ for $l = 0$. In obtaining (3.15), we made use of the fact that (3.12) holds if, and only if, it holds for the log of the terms on each side of the minus sign. The parameters in (3.15), a', is the set of coefficients in a linear regression in which the explanatory variables are all orthogonal. As a result, there is no multicollinearity problem, even if N^p is quite large. For example, we have applied the algorithm with N^p as high as 100. In contrast, we had difficulty executing the regression step in conventional PEA (see (3.10)) for N^p larger than 5, because of multicollinearity problems.[20] Second, the Chebyshev interpolation theorem suggests that it is a good idea to select grid points using the roots of a Chebyshev polynomial (see Judd (1992a,1992b) for a formal statement of the theorem). Suppose we have a given value of a, based on some fixed value of N^p. According to the Chebyshev polynomial approximation theorem, if $N^p \to \infty$ in the computation of a', then

$$\sup_{k \in (\underline{k}, \bar{k}), \theta \in \Theta} \|R(k, \theta; \hat{g}_{a'})\| \to 0 \text{ as } N^p \to \infty, \tag{3.16}$$

when a' is given by (3.15). Thus, the theorem suggests that with large N^p, the function $\exp(\hat{\epsilon}_{a'}(k_i, \theta))$ will not display pathological behavior between grid points.[21] This is an attractive property that is not satisfied by polynomial interpolation schemes generally.

Other weighted residual methods can also be used to find a^*. One such method, Galerkin, is discussed below. We chose to go with collocation because it allows us to convert the nonlinear regression step in the conventional PEA into a linear regression step. For example, this conversion is not possible with Galerkin, which sets weighted averages of R to zero. This impossibility reflects the fact that the log of an average is not equal to the average of the log.

To summarize, the PEA can be viewed as a weighted residual method. In this context, the

[20] den Haan and Marcet (1990) report similar difficulties.

[21] It is important to emphasize what (3.16) does *not* say. Let us make the dependence on a of R in equation (3.12) explicit by writing $R(k, \theta; \hat{g}_{a'}, a)$. Then, (3.16) does not say that $\sup_{k \in (\underline{k}, \bar{k}), \theta \in \Theta} \|R(k, \theta; \hat{g}_{a^*}, a^*)\| \to 0$ as $N^p \to \infty$. We are currently working on a proof of this latter proposition.

three versions of the PEA are differentiated according to the weighting functions used and the manner of evaluating the integrals. The conventional PEA puts relatively heavy weight on (k, θ) pairs with high probability and evaluates all integrals in the analysis using Monte Carlo integration.[22] The PEA with exogenous oversampling shifts more weight into exogenously specified regions, but otherwise pursues the same computational strategy as the conventional PEA. We also described PEA-collocation. This appears to have several advantages relative to conventional PEA: the nonlinear regression step with multicollinear explanatory variables in conventional PEA is transformed into a linear regression with orthogonal explanatory variables; the number of observations in the regression step is very small, and equals the number of parameters in the parameterized expectation function, i.e., $T = 2N^p$, which is no greater than 16 in our experiments (in conventional PEA, T can be in the tens of thousands); and the distribution of (k, θ) is more dispersed, thus ameliorating the conventional PEA's problem that it tends to concentrate observations too much.

3.1.2. Galerkin

Judd (1992a) has discussed approximating policy functions by Chebyshev polynomials and applying the Galerkin method. In our version of this approach, we proceed as follows.[23] The decision rules are:

$$g(k, \theta) \approx \hat{g}_a(k, \theta) \equiv \sum_{i=0}^{N(\theta)-1} a_i(\theta) T_i(\varphi(k)). \tag{3.17}$$

The basis functions for $\hat{g}_a$ are the N^J Chebyshev polynomials. The $[N(\sigma) + N(-\sigma)] \times 1$ vector

$$a = \{a_i(\theta) \mid i = 0, 1, \ldots, N(\theta) - 1, \theta \in \Theta\}$$

[22] In his comment on conventional PEA, Judd (1993) expresses concern about the absence of a solid rationale for sampling at high probability points, or for using Monte Carlo rather than quadrature integration. Our computational results in section 5 below have nothing to say about the latter point, but do suggest that sampling at high probability points is inefficient.

[23] For a case study comparing the method discussed in this subsection with a log-linearization procedure, see Chari, Christiano, and Kehoe (forthcoming).

contains the as-yet-undetermined scalar coefficients and $\varphi(\cdot)$ is defined in (3.6). For now, we suppose that $N(\sigma) = N(-\sigma) = N^J$. The $2N^J$weighting functions, $w(k,\theta)$, are constructed from the basis functions. They are:

$$w^l(k,\theta) = \frac{1}{(1-\varphi(k)^2)^{1/2}} \frac{d\hat{g}_a(k,\theta)}{da_l}, \tag{3.18}$$

for $l = 1,\ldots,2N^J$.

We evaluate (3.2) using $M-$point Gauss-Chebyshev quadrature. To do this, we first compute the $M > N^J$ roots, r_i, $i = 1,\ldots,M$, of the M^{th} order Chebyshev polynomial and use these to construct a grid of capital stocks that is stored in the $M \times 1$ vector $\vec{k}$, $\vec{k} = [\varphi^{-1}(r_1), \varphi^{-1}(r_2), \ldots, \varphi^{-1}(r_M)]'$. Second, we form the $N^J \times M$ matrix A of rank N^J :

$$A = \begin{bmatrix} T_0(r_1) & T_0(r_2) & \cdots & T_0(r_M) \\ T_1(r_1) & T_1(r_2) & \cdots & T_1(r_M) \\ \vdots & \vdots & \vdots & \vdots \\ T_{N^J-1}(r_1) & T_{N^J-1}(r_2) & \cdots & T_{N^J-1}(r_M) \end{bmatrix}. \tag{3.19}$$

Using this notation, the Gauss-Chebyshev quadrature approximation of (3.2) is written compactly, in matrix form, as follows:

$$A\vec{R}(\vec{k},\theta;a) = 0, \theta \in \Theta, \tag{3.20}$$

where $\vec{R}(\vec{k},\theta;a) \equiv [R(k_1,\theta;a), R(k_2,\theta;a), \ldots, R(k_M,\theta;a)]'$. Equation (3.20) represents a nonlinear system of $2N^J$ equations in the $2N^J$ unknowns, which can solved using standard computational routines. Below, we refer to this method as Spectral-Galerkin.

3.2. Two Finite Element, Weighted Residual Methods

We consider the simplest class of finite element functions, those which approximate the policy function with a $\hat{g}_a(k,\theta)$ taken from the space of functions that is piecewise linear and

continuous in k for each fixed θ.[24] The parameters of this function are the values of $k' \equiv \hat{g}$ at each point on a grid of N^f capital stocks, for each θ. Denote this capital grid by a vector K, with elements ordered from smallest to largest, $K = (k_1, k_2, \ldots, k_{N^f})'$. Here, $k_1 \geq \underline{k}$ and $k_{N^f} \leq \overline{k}$. Also, denote the value of k' at each (k_i, θ) by $a_i(\theta)$, for $\theta \in \Theta$, $i = 1, 2, \ldots, N^f$. The $2N^f \times 1$ vector a denotes the set of these parameters. The formal representation of $\hat{g}_a$ is:

$$\hat{g}_a(k, \theta) \equiv \sum_{i=1}^{N^f} a_i(\theta) M_i(k), \tag{3.21}$$

for $\theta \in \Theta$. The basis functions for $\hat{g}_a$ are the N^f functions, $M_i(\cdot)$, which are defined as follows:

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

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

$$M_1(k) = \begin{cases} \frac{k_2 - k}{k_2 - k_1}, & k_1 \leq k \leq k_2 \\ 0, & \text{elsewhere,} \end{cases}$$

and

$$M_{N^f}(k) = \begin{cases} \frac{k - k_{N^f - 1}}{k_{N^f} - k_{N^f - 1}}, & k_{N^f - 1} \leq k \leq k_{N^f} \\ 0, & \text{elsewhere.} \end{cases}$$

After specifying a set of $2N^f$ weighting functions, $w(k, \theta)$, equation (3.2) is used to pin down values for a. An advantage of finite element methods is computational speed. The fact that the parameters of the finite element method have only local impact implies that the number of operations needed to solve this system of equations is smaller by orders of magnitude than is the case in, for example, the spectral methods.[25] In the context of the finite element methods with collocation, existing efforts to realize this computational

[24] 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.

[25] Let n denote the dimension of a. By order of magnitude of the operation count, we mean an integer j such that $c(n)/n^j$ —a non-zero constant as $n \to \infty$, where $c(n)$ is the number of operations needed to compute a. For example, $j = 3$ in the Spectral-Galerkin method because of the matrix inversion involved in applying Newton-Raphson.

efficiency have centered on a particular *time stepping* algorithm (see, for example, Bizer and Judd (1989), Coleman (1988), and Danthine and Donaldson (1981)). In the context of finite element methods with Galerkin, Judd (1991, p. 12) and McGrattan (1993) point out that sparse matrix inversion methods can cut the order of magnitude of the operation count.[26]

3.2.1. Collocation

The finite element, weighted residual method with collocation chooses a so that

$$R(k_i, \theta; \hat{g}_a) = 0, \tag{3.22}$$

for $i = 1, 2, \ldots, N^f$ and $\theta \in \Theta$. This is (3.2), with the weighting functions, $w(k, \theta)$, constructed using dirac-delta functions analogous to those used for Quadrature PEA. Equation (3.22) is a nonlinear system of $2N^f$ equations in $2N^f$ unknowns. Coleman and others apply the following *time stepping* method for solving (3.22):

1. Fix a. (We use starting values based on a log-linear approximation to the model's solution.)

2. For each element of the capital grid k_i find the $k_i'(\theta)$ that solves

$$\begin{aligned} &U_c(k_i, k_i'(\theta), \theta) = \\ &\beta\{(\tfrac{1}{2})U_c(k_i'(\theta), \hat{g}_a(k_i'(\theta), \sigma), \sigma)[f_k(k_i'(\theta); \sigma) + 1 - \delta] \\ &+(\tfrac{1}{2})U_c(k_i'(\theta), \hat{g}_a(k_i'(\theta), -\sigma), -\sigma)[f_k(k_i'(\theta); -\sigma) + 1 - \delta]\}, \end{aligned} \tag{3.23}$$

   for $\theta = -\sigma, \sigma$.

3. Set $a' = \{k_i'(\theta),\ i = 1, 2, \ldots N^f,\ \theta \in \Theta\}$.

4. If the maximum deviation of a' and a exceeds some chosen tolerance, set $a = a'$ and go to step 2. Otherwise, a' is the value of a sought.

[26] In our application of the finite element method with Galerkin, we did not apply a sparse matrix inversion routine.

Below, we refer to this algorithm as FEM-collocation.

3.2.2. Galerkin

The finite element, weighted residual method with Galerkin has been advocated by McGrattan (1993). In our example, the method works to select the value of a that solves the version of (3.2) in which the $2N^J$ weighting functions, $w(k,\theta)$, are constructed from the basis functions, $M_i(k), i = 1,2,\ldots N^J$ in a manner exactly analogous to (3.18) or (3.13)-(3.14). The integral in (3.2) is approximated by a finite sum using $M-$point, Gauss-Legendre quadrature. The algorithm then solves the analog equations to (3.20) by Newton-Raphson methods.[27] Below, we refer to this algorithm as FEM-Galerkin.

4. Solving the Constrained Model

This section describes modifications to the algorithms discussed in the previous section, which are designed to accommodate the irreversible investment version of our model. We pursue two types of modifications. One is based on the Euler equation associated with the Lagrangian representation of the constrained problem. The other is based on the Euler equation associated with the penalty function representation of the problem. We apply the Lagrangian method in the context of both spectral methods and FEM-Galerkin. For reasons elaborated on below, we apply the penalty function method in the context of the FEM only.

The penalty function method is conceptually straightforward. (See Reddy for a discussion.) Since the Lagrangian method is less straightforward, we begin this section by

[27]Taking into account the region over which M_i is zero, equation (3.2) is: $\int_{k_{i-1}}^{k_{i+1}} R(k,\theta,\widehat{g}_a)M_i(k)dk$ for $i = 2,\ldots,N^J-1$ and $\int_{k_i}^{k_{i+1}} R(k,\theta,\widehat{g}_a)M_i(k)dk$ for $i = 1$, and $\int_{k_{i-1}}^{k_i} R(k,\theta,\widehat{g}_a)M_i(k)dk$ for $i = N^J$. These expressions are (3.2) with weighting functions defined analogous to (3.13)-(3.14), but the dirac delta functions replaced by $M_i(k)$. M-point Gauss-Legendre quadrature integration of each integral involves selecting a grid of M capital stocks, say $\vec{k}_i$, over the associated range of integration using the following procedure. First, use the Gauss-Legendre quadrature formulas (see Press, Teukolsky, Vetterling, and Flannery (1992, pp. 140-153)) to select M points in the interval $(-1,1)$. Then, the elements of $\vec{k}_i$ are obtained by applying $\varphi^{-1}(\cdot)$ to these numbers. Second, we compute the MN^J-dimensional vectors $\vec{k} = [\vec{k}_1,\ldots,\vec{k}_{N^J}]$ and $\vec{R}(\vec{k},\theta;a)$. An $N^J \times MN^J$ matrix A is computed so that $A\vec{R}(\vec{k},\theta;a)$ represents the finite sum approximation of the integral in (3.2).

presenting our basic strategy for applying it. Under the Lagrangian method, we seek two functions: one relating the capital decision to the state and the other relating the Lagrange multiplier to the state. It is computationally efficient to restrict the space in which these functions belong. We impose two types of restrictions. The first type is valid generally, and simply enforces the Kuhn-Tucker conditions. The second type of restriction involves assumptions about the properties of the exact functions that solve the problem. We assume that (i) the irreversibility constraint is never binding for the high value of the shock, (ii) the capital policy function is continuous, (iii) the multiplier policy function is continuous, and (iv) for fixed θ, if the constraint is binding from some level of k then it is also binding for all higher levels of k. In practice, the validity of these assumptions can be verified *ex post* by studying the Euler residual function associated with a proposed numerical solution. We do this in section 5, where we report our numerical results. There we also evaluate the validity of our assumptions by studying the solution to our problem based on a dynamic programming algorithm.

Figure 1 depicts hypothetical policy functions in $k \times k'$ space. The policy functions are drawn for a case in which gross investment has *not* been constrained to be nonnegative, and in fact does take on negative values for some values of k. The points at which the $g(k,\sigma)$ and $g(k,-\sigma)$ functions cross the 45° line mark the ergodic set of capital for this case. Notice that $g(k,\sigma)$ never crosses the $(1-\delta)k$ line within the ergodic set. This implies that when $\theta = \sigma$ investment is never negative. If, when we impose the nonnegativity constraint, $g(k,\sigma)$ retains its basic shape, then we can infer that λ will always be zero when $\theta = \sigma$. On the other hand, we see from the figure that $g(k,-\sigma)$ crosses the $(1-\delta)k$ line at a point within the ergodic set. Based on this result, we conjecture that when we impose the nonnegativity constraint, the exact policy function has the following property: when $k \geq \tilde{k}$, for some $\tilde{k}$, and $\theta = -\sigma$ the nonnegativity constraint binds and $\lambda \geq 0$, and when $k < \tilde{k}$ and $\theta = -\sigma$ the nonnegativity constraint does not bind and $\lambda = 0$.

We use these observations and the requirement that the Kuhn-Tucker conditions be satisfied to constrain the space of approximate policy functions. Our task is to find functions

to approximate two policy functions: $g(\cdot)$ as before, and the function determining λ, $h(\cdot)$. We restrict the space of approximating functions for $g(\cdot)$ as follows:

$$g(k,\sigma) \approx \hat{g}_a(k,\sigma), \tag{4.1}$$

$$g(k,-\sigma) \approx \hat{g}_a(k,-\sigma) \equiv \begin{cases} \max\{\tilde{g}_a(k),(1-\delta)k\}, & k < \tilde{k} \\ (1-\delta)k, & \tilde{k} \leq k \end{cases}, \tag{4.2}$$

where $\tilde{k}$ is a function of a and is defined by the property

$$\tilde{g}_a(\tilde{k}) = (1-\delta)\tilde{k}. \tag{4.3}$$

We restrict the space of approximating functions for $h(\cdot)$ as follows:

$$h(k,\sigma) \approx \hat{h}_a(k,\sigma) = 0, \tag{4.4}$$

and

$$h(k,-\sigma) \approx \hat{h}_a(k,-\sigma) \equiv \begin{cases} 0, & k < \tilde{k} \\ \max\{\tilde{h}_a(k),0\}, & \tilde{k} \leq k \end{cases}, \tag{4.5}$$

with the property

$$\tilde{h}(\tilde{k}) = 0. \tag{4.6}$$

The first expression just says that we use some convenient function $\hat{g}_a(\cdot,\sigma)$, *i.e.* a polynomial or a piecewise linear function parameterized by the vector a, to approximate the true rule $g(\cdot,\sigma)$. Expression (4.2) embodies the restriction that the approximating function for $g(\cdot,-\sigma)$, $\hat{g}_a(\cdot,-\sigma)$, must force gross investment to be zero for $k \geq \tilde{k}$. A polynomial or a piecewise linear function $\tilde{g}_a(\cdot)$ is used to approximate the true policy function for $k < \tilde{k}$. The max operator in (4.2) ensures the constraint on non-negative gross investment is never violated. Also, by solving for $\tilde{k}$ using (4.3) we make sure $\hat{g}_a(\cdot,-\sigma)$ is continuous. In (4.4) we use our conjecture that the gross investment constraint never binds when $\theta = \sigma$ to force λ to be zero for all k in this case. For the case $\theta = -\sigma$ we force λ to be zero for $k < \tilde{k}$ and we propose a

polynomial or piecewise linear function $\tilde{h}_a(\cdot)$ be used to approximate $h(\cdot, -\sigma)$ for $k \geq \tilde{k}$. The max operator is used to ensure the Lagrange multiplier is always non-negative and condition (4.6) forces the function that approximates $h(\cdot, -\sigma)$, $\hat{h}_a(\cdot, -\sigma)$, to be continuous.

It should be clear from expressions (4.1)-(4.6) that our space of approximating functions forces the Kuhn-Tucker conditions to be satisfied exactly. Our only task is to find approximating functions within this class that set the Euler residuals to zero for admissible k and θ. This task is not unlike the one we encountered when we ignored constraint (2.2), which suggests we can apply similar methods to solve the constrained model.

Assumptions (i)-(iv), stated at the beginning of this section, play an important role in our Lagrange multiplier procedure. In section 5 below, we report evidence that these assumptions are valid in our model, so that imposing them is innocuous in our application. However, they may be difficult to impose in higher dimensions, or there may be models in which the assumptions are not valid. Because of this, it is useful to note that there exist versions of our Lagrange multiplier method which do not involve assumptions (i)-(iv). For example, we can modify our procedure so that (iii)-(iv) are dropped by replacing (4.2)-(4.3) by $g(k, -\sigma) \approx \hat{g}_a(k, -\sigma) = \max\{\tilde{g}_a(k), (1-\delta)k\}$ for all $k > 0$, and (4.5)-(4.6) by

$$h(k, -\sigma) \approx \hat{h}_a(k, -\sigma) \equiv \begin{cases} 0, & \hat{g}_a(k, -\sigma) > (1-\delta)k \\ \max\{\tilde{h}_a(k), 0\}, & \hat{g}_a(k, -\sigma) \leq (1-\delta)k \end{cases}.$$

This perturbation on our method does not involve computing the variable, $\tilde{k}$. We have conducted several experiments with this procedure and found it to be practical.

We now turn to the description of our modifications to the algorithms discussed in the previous section.

4.1. A Lagrangian Modification to the PEA

The modification to the PEA to accommodate the case where (2.2) binds occasionally is remarkably straightforward. We must now be careful to allow the function m in (2.6) to accommodate a potentially nonzero multiplier, h. To do this, we work with the following

modified version of (3.3):

$$\int_{\theta'} m(k,\theta,\theta';g,h)p_{\theta'}(\theta' \mid \theta)d\theta' \approx \exp[\hat{e}_a(k,\theta)], \tag{4.7}$$

where $\hat{e}_a(k,\theta)$ is defined in (3.5) and m is defined in (2.6). The PEA's approximation to the decision rule is:

$$\hat{g}_a(k,\theta) = \max\left\{(1-\delta)k, \exp(\theta)k^\alpha + (1-\delta)k - U_c^{-1}\left[\beta \exp\left(\hat{e}_a(k,\theta)\right)\right]\right\}, \tag{4.8}$$

and its approximation to the multiplier function, $\hat{h}_a(k,\theta)$, is:

$$\hat{h}_a(k,\theta) = U_c\left(k, \hat{g}_a(k,\theta), \theta\right) - \beta \exp[\hat{e}_a(k,\theta)]. \tag{4.9}$$

With these modifications to $\hat{g}_a$ and $\hat{h}_a$, the three versions of the PEA can be implemented as described in the previous section.

4.2. A Lagrangian Modification to the Spectral-Galerkin Method

We choose functional forms for $\hat{g}_a(\cdot,\sigma)$, $\tilde{g}_a(\cdot)$, and $\tilde{h}_a(\cdot)$ as follows:

$$\hat{g}_a(k,\sigma) = \sum_{i=0}^{N(\sigma)-1} a_i(\sigma)T_i(\varphi(k)), \tag{4.10}$$

$$\tilde{g}_a(k) = \sum_{i=0}^{N(-\sigma)-1} a_i(-\sigma)T_i(\varphi(k)), \tag{4.11}$$

and

$$\tilde{h}_a(k) = \sum_{i=0}^{N^\lambda} b_i T_i(\varphi(k)). \tag{4.12}$$

An Euler residual function can be constructed in the manner used before to form $R(k,\theta;a)$, where a is the $[N(\sigma)+N(-\sigma)+N^\lambda]\times 1$ vector of unknown polynomial coefficients.

To apply the Spectral-Galerkin method we must find a grid vector $\vec{k}$ and weighting matrix A that can be used to form the system $A\vec{R}(\vec{k},\theta;a)$. We can use a Chebyshev grid as before

to construct $\vec{k}$. When constructing the grid we made sure it was fine enough so that there were a substantial number of grid points to the right of $\tilde{k}$. Presumably this is needed to ensure a good approximation to $h(\cdot)$.

We set $N(\sigma) = N(-\sigma) + N^{\lambda} = N^{J}$, $N^{\lambda} = N(-\sigma) - 1$, and selected M, the number of elements in $\vec{k}$, such that $M > N^{J}$. We construct the matrix A as shown in (3.19). The approximation problem is then identical to the one described before: find a that solves $A\vec{R}(\vec{k}, \theta; a) = 0$, for $\theta \in \Theta$.[28]

Our Lagrangian modification of Spectral-Galerkin accommodates nondifferentiable decision rules. This seems appropriate in problems with occasionally binding constraints. We found it less convenient to accommodate nondifferentiable decision rules in the context of a penalty function version of Spectral-Galerkin, so we did not pursue this.

4.3. A Lagrangian Modification to the FEM-Collocation Method

This section describes how we applied the Lagrangian method to FEM-collocation.[29] We choose piecewise linear functions to form $\hat{g}_a(\cdot, \sigma)$, $\tilde{g}_a(\cdot)$, and $\tilde{h}_a(\cdot)$ and select the capital stock grid $\vec{k} = (k_1, k_2, \ldots, k_{N^J})'$. The objective is to solve for the coefficients associated with this grid: $a_i(\theta)$, $i = 1, 2, \ldots, N^J$, $\theta \in \Theta$, as before, and b_i, $i = 1, 2, \ldots, N^J$, where each b_i corresponds to the value taken by the Lagrange multiplier at the i'th element of $\vec{k}$ when $\theta = -\sigma$. Stack the undetermined coefficients in the vector

$$a = (a_1(\sigma), a_2(\sigma), \ldots, a_{N^J}(\sigma), a_1(-\sigma), a_2(-\sigma), \ldots, a_{N^J}(-\sigma), b_1, b_2, \ldots, b_{N^J})'.$$

[28] If $N(\sigma) \neq N(-\sigma) + N^{\lambda}$, we can use the $M \times 1$ Chebyshev grid as in the previous case, but now we must choose separate weighting matrices for the Euler residual functions $R(\vec{k}, \sigma; a)$ and $R(\vec{k}, -\sigma; a)$. These weighting matrices, call them A^h and A^l, can be constructed in a manner analogous to the construction of A in the previous case. The approximation problem is to find a that solves $A^h\vec{R}(\vec{k}, \sigma; a) = A^l\vec{R}(\vec{k}, -\sigma; a) = 0$.

[29] See Coleman, Gilles, and Labadie (1992) for another application of the Lagrangian method in the context of the FEM-collocation algorithm.

We modify the FEM-collocation algorithm as follows. In step 2 of that algorithm equation (3.23) for $\theta = \sigma$ is replaced by

$$\begin{aligned} &U_c(k_i, k_i'(\sigma), \sigma) = \\ &\beta\{(\tfrac{1}{2})U_c(k_i'(\sigma), \widehat{g}_a(k_i'(\sigma), \sigma), \sigma)[f_k(k_i'(\sigma), \sigma) + 1 - \delta] \\ &+(\tfrac{1}{2})(U_c(k_i'(\sigma), \widehat{g}_a(k_i'(\sigma), -\sigma), -\sigma)[f_k(k_i'(\sigma), -\sigma) + 1 - \delta] \\ &-\widehat{h}_a(k_i'(\sigma), -\sigma)(1-\delta))\}. \end{aligned} \tag{4.13}$$

Equation (3.23) for $\theta = -\sigma$ is replaced by:

$$\begin{aligned} &U_c(k_i, k_i'(-\sigma), -\sigma) - \lambda_i = \\ &\beta\{(\tfrac{1}{2})U_c(k_i'(-\sigma), \widehat{g}_a(k_i'(-\sigma), \sigma), \sigma)[f_k(k_i'(-\sigma), \sigma) + 1 - \delta] \\ &+(\tfrac{1}{2})(U_c(k_i'(-\sigma), \widehat{g}_a(k_i'(-\sigma), -\sigma), -\sigma)[f_k(k_i'(-\sigma), -\sigma) + 1 - \delta] \\ &-\widehat{h}_a(k_i'(-\sigma), -\sigma)(1-\delta))\}. \end{aligned} \tag{4.14}$$

For each i equation (4.13) is solved by choice of $k_i'(\sigma)$ as before. Equation (4.14) is first solved by choice of $k_i'(-\sigma)$ with $\lambda_i = 0$. If $k_i'(-\sigma) > (1-\delta)k_i$ then we proceed to the next value of i in the sequence $i = 1, 2, \ldots, N^f$. Otherwise, $k_i'(-\sigma)$ is set equal to $(1-\delta)k_i$ and (4.14) is solved by choice of λ_i. The only other modification to the algorithm is to add to the updating rules of step 3 the conditions $b_i = \lambda_i$, $i = 1, 2, \ldots, N^f$.[30]

4.4. A Penalty Function Modification to the FEM-Galerkin Method

We now turn to the penalty function implementation of FEM-Galerkin. This is a modified version of the algorithm applied by McGrattan (1993). In this approach, a penalty is applied to violations of the constraint on capital accumulation. Specifically, we solve a modification

[30]There are two technical matters to be resolved regarding the implementation of side conditions (4.3) and (4.6). The function $\widetilde{g}$ is defined as a set of linear segments which are joined at $\widetilde{g}_a(k_i) \equiv a_i(-\sigma)$ for $i = 1, 2, \ldots, N^f$ and at $\widetilde{g}_a(\widetilde{k}) = (1-\delta)\widetilde{k}$, where $\widetilde{k}$ is obtained as follows. Examine $a_i(-\sigma)$ for $i = 1, 2, \ldots, N^f$ until the first i, say i' occurs with $a_i(-\sigma) \leq (1-\delta)k_i$. Then use the line segment defined by $\widetilde{g}_a(k_{i'-2})$ and $\widetilde{g}_a(k_{i'-1})$ to linearly extrapolate a value for $\widetilde{k}$. Formally, $\widetilde{g}$ is defined exactly as $\widehat{g}$ is in (3.21), with the exception that $\widetilde{k}$ is added into the list $k_1, k_2, \ldots, k_{N^f}$. We impose (4.6) by defining $\widetilde{h}$ to be composed of linear segments joined at $\widetilde{h}_a(\widetilde{k})$ and at $\widetilde{h}_a(k_i) \equiv b_i$ for $i = 1, 2, \ldots, N^f$, with $\widetilde{h}_a(\widetilde{k}) \equiv 0$.

to the original planner's problem as follows:

$$W(k,\theta) = \max_{k' \geq 0} \; U(c(k,k',\theta)) + \beta \int_{\theta'} W(k',\theta') p_{\theta'}(\theta' \mid \theta) d\theta' - \frac{\pi}{2} [\max\{0, (1-\delta)k - k'\}]^2 . \tag{4.15}$$

Here π is a nonnegative penalty parameter. For $\pi = 0$, (4.15) describes the problem for the model when the gross investment constraint is ignored. For positive π, violations of (2.2) reduce the planner's objective function. Intuitively, we might expect that for large values of π the solution to problem (4.15) would be "close" to the exact solution of the constrained problem.[31]

We apply the penalty function method by solving the sequence of problems corresponding to an increasing sequence $\{\pi_n\}$. In a typical experiment, the sequence contained 31 elements beginning with 1, 2, 10, 20, 50, ... and ending with 1,200. For each value of π_n, it is necessary to solve, using the FEM-Galerkin method, the Euler equation associated with (4.15):

$$\begin{aligned} &U_c(k, g(k,\theta), \theta) - \pi_n \max\{0, (1-\delta)k - g(k,\theta)\} \\ &-\beta \textstyle\int_{\theta'} \{U_c(g(k,\theta), g(g(k,\theta),\theta'), \theta')[f_k(g(k,\theta),\theta') + (1-\delta)] \\ &-(1-\delta)\pi_n \max[0, (1-\delta)g(k,\theta) - g\,(g(k,\theta),\theta')]\} p_{\theta'}(\theta' \mid \theta) d\theta' = 0. \end{aligned}$$

The algorithm stops when the maximum violation of the gross investment constraint on the capital stock grid is smaller than some prespecified tolerance.[32] Denote by π^* the value of the penalty parameter when the algorithm is completed. Then, following Luenberger (1969,

[31] Luenberger (1969, Theorem 1, p. 306) provides a theorem for the case where the solution to a constrained maximization problem is a finite dimensional vector. In this case, solutions to the penalty function version of the problem corresponding to an increasing sequence of penalty parameters tending toward infinity will converge to the exact solution. Presumably it is straightforward to extend the theorem to our environment.

[32] To be precise, let

$$\gamma(\pi_n) = \max_{i \in \{1,2,\ldots,NJ\}, \theta \in \Theta} \{\max[0, (1-\delta)k_i - \hat{g}_a(k_i,\theta)]\}.$$

Then, the algorithm stops when a value of π_n is encountered with $\gamma(\pi_n)$ less than the chosen tolerance. Solving a sequence of models corresponding to an increasing sequence $\{\pi_n\}$ is computer and programmer-time intensive. When we instead attempted to initiate the algorithm with a large value of π, however, the algorithm crashed. This was because our initial guess at the solution (namely, the log-linear approximation) was not a good one in this case.

Theorem 2, p. 307), an approximation to $h(k,\theta)$ is given by:

$$\widehat{h}_a(k,\theta) = \pi^* \max\{0, (1-\delta)k - \widehat{g}_a(k,\theta)\}.$$

It is worth noting that the FEM is particularly suited to working with penalty functions. This is because it uses a functional form that easily accommodates nondifferentiabilities in the exact policy function. As we shall see below, there is a likelihood that the solution to the constrained problem will involve a policy function for capital that is not smooth. If we were to apply the Spectral-Galerkin method with a penalty function we would encounter difficulties because that method attempts to approximate the exact policy function with a globally smooth approximator.

5. Evaluating the Algorithms

The algorithms we have described were used to approximate the solution to a particular parameterization of the model. In addition to applying these algorithms we also approximated the model solution using dynamic programming (DP) applied to a discrete version of the model. We take the DP solution to be virtually identical to the exact solution and use it as a benchmark for evaluating the algorithms discussed in this paper. Details about these computations are reported in appendix 1. One of our findings is that the results of all the algorithms are reasonably accurate. The differences between solution methods are small and not economically very meaningful in the context of our model economy. Nonetheless, there are some noticeable differences in accuracy and in computation time, and we think these are potentially useful as input into decisions about which algorithm to use in more complex modeling environments.

We study three aspects of the approximate solutions: the Euler residuals, the policy functions, and the implications of the policy functions for various first and second moment properties of the model. With one exception, the parameter values we chose are standard in the real business cycle literature and are as follows: $\alpha = 0.3$, $\delta = 0.02$, $\beta = 1.03^{-.25}$, and

$\sigma = 0.22$. The exceptional case, σ, was chosen large enough to ensure that the investment constraint binds occasionally. Finally, we specified $U(c_t) = \ln c_t$.

We study three cases for the PEA and two cases for each of the other algorithms. For the PEA, $N^p = 3$ in all cases except PEA-collocation applied to the irreversible investment model, in which case $N^p = 8$. The solution labelled $N^p = 3$ corresponds to what we have called conventional PEA, while the one labelled $N^p = 3^*$ is PEA with exogenous oversampling starting near an estimate of $\tilde{k}$. The solutions labelled $N^p = 3^{**}$ and $N^p = 8^{**}$ correspond to PEA-collocation applied in the reversible and irreversible investment cases, respectively.[33] For the Spectral-Galerkin algorithm the cases are: $N^J = 3$ and 8 in the reversible investment case and $(N(\sigma), N(-\sigma), N^\lambda) = (5,3,2)$ and $(9,5,4)$ in the irreversible investment case. For the FEM-collocation algorithm the two cases are: $N^J = 36$ and 72. In the case of the FEM-Galerkin algorithm we study solutions based on $N^J = 18$ and 36.[34] For each algorithm, the highest order parameterization reported is the one for which we obtained convergence. Our convergence criterion was based on the second moment properties reported in tables 2 and 3. For each method we incremented the number of parameters until the change in all second moment properties was less than one standard deviation. Typically, the last moments to converge were the ones based on financial variables.

All computations were carried out on a Gateway 2000 486 DX2/66 and the programs

[33]For the conventional and modified PEA algorithms, we set $T = 10,000$. This compares to a value of $T = 2,500$ used by den Haan and Marcet (1990). They work with a model similar to ours and assume the technology shock standard deviation is 0.32, which contrasts with a standard deviation of 0.22 in our model (the one-step-ahead conditional standard deviation in the technology shock in their model is 0.10). For conventional PEA with oversampling, we set $T = 7,500$, $J = 100$, $\tilde{T} = 25$, and k_0^j corresponding to a number (34.0) in the neighborhood of $\tilde{k}$, for $j = 1, \ldots, J$.

[34]Additional details of how the algorithms were implemented are as follows. For the Spectral-Galerkin cases we used $M = 100$. For the parameterization of the model we examined, this guarantees an ample supply of grid points on either side of $\tilde{k}$. The grids for the FEM algorithms were chosen to be equally spaced between boundaries just outside the initial guess for the ergodic set, $(\underline{k}, \bar{k})$. The tolerance on violations of the investment constraint for the penalty function version of the FEM-Galerkin algorithm was set at 5×10^{-5}. This tolerance was reached for $\pi^* = 1,250$ with $N^J = 18$ and $\pi^* = 1,200$ with $N^J = 36$. In the text, the approximate policy functions were expressed in terms of the level of the capital stock. We did this to simplify the presentation. In the calculations we work in terms of the log of the capital stock. Grids for the Spectral-Galerkin and PEA-collocation methods were constructed based on the log of the capital stock, while grids for the FEM were constructed based on the level of the capital stock.

are available on request. The computation times are displayed in table 1.[35] These times should be interpreted with caution. First, we did not make extensive efforts to optimize the computer code. Second, it could be misleading to extrapolate the relative time requirements reported in our experiments to larger problems. For example, the technology shock in our model can take on two values only. This biases computational times in favor of methods such as Spectral-Galerkin and PEA-collocation which exploit this fact, and against conventional PEA which does not. Also, the operation counts of the various algorithms are of different orders of magnitude. For example, the Spectral-Galerkin algorithm involves a number of operations that grows at the rate of the cube of the number of decision rule parameters sought, while the operation count for the FEM grows linearly or with the square.[36]

Still, there are several observations worth making about the timing of the various algorithms. First, note that PEA-collocation is faster by orders of magnitude than all the other algorithms. Second, in contrast with the other algorithms, the computational time for the PEA's does not increase substantially when the irreversible investment constraint is imposed. This reflects the fact that, in contrast to the other algorithms, taking account of this constraint adds virtually no computational burden to a PEA. Third, the slowest algorithm applied to the irreversible investment model is FEM-Galerkin. This reflects the fact that this algorithm involves repeatedly solving the model for higher values of the penalty function parameter.

[35]With three exceptions, all the programming was done in GAUSS. Two exceptions were the conventional PEA and PEA with exogenous oversampling, which combined FORTRAN with GAUSS. The simulation part of the PEA was programmed in FORTRAN and imported into a GAUSS shell. FORTRAN programming could have reduced the computation times for the FEM-collocation algorithm significantly and the computation times for the FEM-Galerkin algorithm marginally. The second exception is the DP calculations which were done in FORTRAN. The DP algorithms required hours to achieve convergence, but the exact times are not reported on the table.

[36]Christiano and Fitzgerald (1991) elaborate on this observation and illustrate it in a comparison of Spectral-Galerkin and FEM-collocation. It should be noted that to achieve a given degree of accuracy in problems with smooth decision rules, fewer parameters may be required in the Spectral-Galerkin procedure than in the FEM. Also, the assertion about the rate of growth of the operation count in Spectral-Galerkin assumes that a standard Newton-Raphson equation-solving method is used. The operation count for Spectral-Galerkin could be reduced if a more sophisticated version of this algorithm were used.

5.1. Euler Equation Residuals

Here, we focus on the Euler residual function (ERF), defined by $R(k,\theta;a^*)$, where a^* denotes the solved value of a. We consider the ERF's for both the reversible and irreversible investment versions of the model. We study the graphs of the residual functions (see figures 2a and 2b) and the maximum absolute value (MAV) of $R(k,\theta;a^*)$ over $k \in (\underline{k},\overline{k})$, for $\theta = \sigma$ and $-\sigma$, respectively (see table 1). We found it useful, in the context of the PEA, to also compute MAV's over a narrower interval, containing 90% of the simulated capital stocks. The upper and lower boundaries of this set are indicated by vertical lines in the PEA component of figure 2a. For the PEA, MAV numbers not in parentheses in table 1 are based on this 90% confidence interval, while numbers in parentheses are based on the entire interval, $(\underline{k},\overline{k})$. In figures 2-5, results are displayed for a range of capital stocks corresponding roughly to the set, $(\underline{k},\overline{k}) = (22.0, 40.0)$.

We now consider the two left hand columns of figures 2a and 2b and the top panel of table 1, which pertain to the ERF's of the reversible investment model. We begin by summarizing our findings for the PEA. We found that increasing N^p beyond $N^p = 3$ in the context of conventional PEA has relatively little impact on the results. In the interests of saving space, we do not document this finding here. Essentially, results based on conventional PEA converged at $N^p = 2$ and roughly correspond to our $N^p = 3$ findings.[37] Note from figure 2a that the ERF's for conventional PEA and $\theta = \sigma$ are consistently negative over the 90% region of capital stocks. This is a sign of inaccuracy in the conventional PEA's solution.[38] One way to accommodate this is to increase the length of the Monte Carlo simulations, T. However, we found that impractically large values for T are required to achieve a significant degree of improvement in accuracy.[39]

[37]Increasing the order of approximation to $N^p = 5$ does not change the results significantly. We found it difficult to increase N^p above 5.

[38]At the same time, the deviations from zero in the euler errors are not large by one economic measure. The percent increase in consumption which would move current marginal utility of consumption down enough to close a given gap, R, in the euler error is $100\beta cR/(1-\beta cR)$, where c is the level of consumption. The worst MAV in the 90% confidence region for the PEA is the value of 0.00013 reported for conventional PEA, $\theta = -\sigma$. In consumption units this is about 0.03 percent.

[39]We did the $N^p = 3$ calculations for $T = 20,000$, $40,000$, $60,000$, and $80,000$. For $\theta = \sigma$, the MAV's

An alternative to increasing T is to alter the distribution of (k, θ) points at which the computations are done. This can be seen by noting the significant improvements that are obtained in the $N^p = 3^*$ and $N^p = 3^{**}$ versions of the PEA. In particular, note that the ERF's for $\theta = \sigma$ are closer to zero when $N^p = 3^*$ or 3^{**}, and by orders of magnitude in the latter case. We infer from these results that conventional PEA's Monte Carlo procedure for selecting points in the state space for the computations is not optimal. The alternative procedures, PEA with exogenous oversampling and PEA-collocation, seem to work better. In each case the distribution of (k, θ) values used in the calculations is more diffuse relative to that in conventional PEA. We conjecture that this is the basic reason that they do better. The idea is that they do better for the same reason that regression coefficients are more precisely estimated, the greater is the dispersion in the explanatory variables.

Notice from figure 2 that the performance of the conventional PEA and PEA with exogenous oversampling deteriorates significantly in the outer 5% tail areas of the interval $(\underline{k}, \overline{k})$. The dramatic improvement evident with PEA-collocation over the entire interval $(\underline{k}, \overline{k})$ is quite striking in comparison (visually, it is hard to distinguish from zero in the figure), especially since this improvement is achieved by requiring only that the Euler residuals be zero at three points in this interval.

For the other three solution algorithms, increasing the number of parameters in the decision rule is very effective at driving the ERF's toward zero. In each case, convergence to zero is roughly uniform over the range $(\underline{k}, \overline{k})$. Note how smooth the ERF's corresponding to Spectral-Galerkin are, in contrast to those based on the two FEM methods. This reflects the fact that, in our example, the smoothness in the Spectral decision rule mimics more closely the properties of the exact decision rule.

We now consider the two right hand columns of figure 2 and the bottom panel of table 1, which pertain to the ERF's of the irreversible investment model. We again begin by sum-

for these cases are 9.5×10^{-5} (9.5×10^{-5}), 7.4×10^{-5} (3.0×10^{-4}), 5.3×10^{-5} (4.5×10^{-4}), 3.2×10^{-5} (3.4×10^{-4}), respectively. For $\theta = -\sigma$ the MAV's are 1.0×10^{-4} (1.0×10^{-3}), 2.6×10^{-5}, (1.2×10^{-4}), 2.3×10^{-5} (1.5×10^{-4}), 3.5×10^{-5} (9.7×10^{-5}). As in table 1, numbers not in parentheses correspond to MAV's based on an interval containing 90% of the realizations of the capital stock. Numbers in parentheses correspond to MAV's based on an interval that contains 100% of the realizations.

marizing our findings for the PEAs. The performance of the $N^P = 3$ and 3^* versions of this algorithm is roughly comparable to what it is in the reversible investment case. In particular, we found that increasing N^p does not contribute much to accuracy in the conventional PEA, but PEA with exogenous oversampling does help.[40] The PEA-collocation ERFs deteriorate somewhat (particularly around the point at which the investment constraint begins to bind: $k = \tilde{k} = 33.40$, $\theta = -\sigma$, according to DP), but still dominate the other implementations of the PEA.

We now summarize our results for the other algorithms. Relative to the reversible investment case, the two Galerkin methods have a harder time driving the ERF's zero. The MAV for the best Spectral-Galerkin solution (i.e., $(9, 5, 4)$) is approximately 9×10^{-5}, as opposed to 4×10^{-7} in the reversible investment case.[41] Similarly, the MAV for the best FEM-Galerkin method is 4×10^{-4}, as opposed to 7×10^{-5} in the reversible investment case. By contrast, the rate of convergence for FEM-collocation is comparable across the reversible and irreversible investment models. The reason the Galerkin methods have problems is similar to that emphasized in the case of PEA-collocation. It has to do with the difficulty they have in driving the Euler residuals to zero in the neighborhood of $k = \tilde{k}$.

5.2. Policy Function Comparisons

We now compare the policy function approximations obtained for the two versions of the model and for the four solution algorithms. In figure 3, policy and multiplier functions based on the highest order Spectral-Galerkin method are compared with those based on the DP

[40] We did the $N^p = 3$ calculations in the irreversible investment model for $T = 20,000$, $40,000$, $60,000$, and $80,000$. For $\theta = \sigma$, the MAV's for these cases are 1.1×10^{-4} (2.3×10^{-4}), 8.5×10^{-5} (5.2×10^{-4}), 6.1×10^{-5} (6.6×10^{-4}), and 3.9×10^{-5} (5.4×10^{-4}), respectively. For $\theta = -\sigma$ the MAV's are 1.2×10^{-4} (2.0×10^{-4}), 8.2×10^{-5} (7.8×10^{-4}), 8.2×10^{-5} (7.2×10^{-4}), and 9.4×10^{-5} (7.9×10^{-4}). As in table 1, numbers not in parentheses correspond to MAV's based on an interval containing 90% of the realizations of the capital stock. Numbers in parentheses correspond to MAV's based on an interval that contains 100% of the realizations.

[41] We considered higher degree approximations for the Spectral-Galerkin method but were unable to achieve anything resembling the convergence to zero noted in the reversible investment case. For example, with $N^J = 8$ in the reversible investment case, the MAVs are about 2×10^{-11} and 7×10^{-11} for $\theta = \sigma$ and $\theta = -\sigma$, respectively, while in the irreversible investment case with $(15, 8, 7)$ the MAVs are about 4×10^{-5} and 5×10^{-5} for $\theta = \sigma$ and $\theta = -\sigma$, respectively

method, for both reversible and irreversible investment versions of the model.

There are three main results in figure 3. First, there is very little difference between the solution based on Spectral-Galerkin and DP. Over most of the range of k the functions are identical to the eye. We infer from this that the Spectral-Galerkin method provides a highly accurate approximation to the solution. Second, the shape of the policy and multiplier functions validate the four assumptions we made when constructing the space of approximating functions in section 4. Third, for $k < \tilde{k}$ the constrained and unconstrained policy functions are virtually identical, while investment is (slightly) lower in the constrained economy with $\theta = \sigma$, when $k > \tilde{k}$. Presumably, this reflects in part a rate of return effect: the payoff from capital investment is lower in the irreversible investment economy, since there is some chance that $\theta' = -\sigma$, in which case the limitation against consuming capital is binding. The net effect of the irreversibility constraint on the average capital stock is quite small, since the impact on investment of the irreversibility constraint is positive when $k > \tilde{k}$ and $\theta = -\sigma$.[42]

The FEM decision rules are indistinguishable from Spectral-Galerkin, and so we do not graph them. It is worth comparing the PEA and Spectral-Galerkin rules, however. Figure 4 compares the PEA and highest order approximation Spectral-Galerkin investment policy functions for the reversible investment economy. Qualitatively, the findings here are consistent with our analysis of the Euler residual functions. First, the PEA with exogenous oversampling appears to do better in the 90% confidence region for capital, than conventional PEA. Also, the greatest inaccuracy in the conventional PEA and PEA with exogenous oversampling appears to be in the lower tail of the capital stock distribution. Finally, the PEA-collocation policy functions appears to be indistinguishable from the Spectral-Galerkin rule.

Figure 5 compares the PEA and the highest order approximation Spectral-Galerkin investment policy functions for the irreversible investment economy. Again, the results here are consistent with our analysis of the Euler residuals in figure 4. Thus, the PEA-collocation

[42]We found that the average capital stock in the reversible and irreversible versions of the model is 31.3 in each case. Uncertainty per se does seem to have an impact since the steady state capital stock is 30.5.

decision rule appears to be essentially indistinguishable from what we take to be more or less the exact decision rule. Also, the decision rule produced by the PEA with exogenous oversampling represents a definite improvement over conventional PEA. One way to see this is that conventional PEA seems to more seriously miss identify $\widetilde{k}$ than PEA with exogenous oversampling. This can be seen most clearly in the graphs of the approximate Lagrange multipliers, computed using (4.9).[43]

5.3. Approximate Model Implied Moments

We now examine several model moments computed using our four approximate policy functions and the DP algorithm. In table 2 moments related to the unconstrained model are displayed. The moments we computed for this case are as follows: R^e (the mean value of $f'(k_t, \theta_t) + (1 - \delta)$), R^f (the mean return on a one-period-ahead state-uncontingent consumption loan), $R^e - R^f$ (the mean equity premium), σ_j, $j = y, c, i$ (the standard deviation of gross output, consumption and gross investment, respectively), $\rho(y, j)$, $j = c, i$ (correlation of gross output with consumption and gross investment, respectively), and freq$(i < 0)$ (the frequency of times that gross investment is negative). The rate of return variables, R^f and R^e, are expressed in annual percent terms. In table 3 moments related to the constrained model are displayed. In addition to the moments displayed in table 2 we compute moments related to Tobin's q, the price of new capital in terms of consumption goods. We define this price as follows: $q = 1 - \lambda / U'(c)$.[44] Also, we replace freq$(i < 0)$ with freq$(\lambda > 0)$ (the frequency of times that the gross investment constraint binds) in table 3. All statistics are based on averages from samples of length 114 replicated 500 times. Numbers in parentheses are Monte Carlo standard errors.[45]

[43]With $N^p = 3$, calculations using conventional PEA for T =10,000, 20,000, 40,000, 60,000, and 80,000, resulted in the following estimates of $\widetilde{k}$: 33.03, 33.15, 33.27, 33.31, 33.34. With $N^p = 3^*$ (i.e., PEA with exogenous oversampling) and $T = 10,000$, we obtained $\widetilde{k} = 33.29$. With $N^p = 8^{**}$ (i.e., PEA-collocation) we obtained $\widetilde{k} = 33.37$. We take the $\widetilde{k}$ implied by DP, which is $\widetilde{k} = 33.40$, to be the exact solution. Thus, a given level of accuracy (in terms of $\widetilde{k}$) can be achieved with a lower value of T by applying alternatives to conventional PEA.

[44]See Sargent (1980) for an analysis of Tobin's q in a setting similar to ours.

[45]These are $\sigma_x / \sqrt{500}$, where σ_x is the standard deviation, across our 500 replications, of some statistic, x.

Consider table 2. With conventional PEA there is slight inaccuracy in its predictions for financial variables (*e.g.* the equity premium is 0.089 % (0.011) with conventional PEA versus 0.049 % (0.010) with DP, with standard errors in parentheses) and the algorithm overestimates the frequency that gross investment is negative (9.89 % (0.38) with conventional PEA and 8.89 % (0.36) with DP.) PEA with exogenous oversampling and PEA-collocation both show improvement along these dimensions. Spectral-Galerkin shows convergence at $N^J = 5$, FEM collocation shows convergence at $N^f = 72$, and FEM Galerkin at $N^f = 36$. Notice FEM collocation with $N^f = 36$ is unable to achieve convergence while FEM Galerkin, using exactly the same approximating function, does achieve convergence. This reflects the relative computational efficiency of smoothly weighting the Euler residuals, which Galerkin does, in the context of the FEM.

Now consider table 3. The primary differences between methods, which in any case are small, lie in their implications for statistics involving financial variables. For example, the main difference between conventional PEA and the other versions of PEA is that the former over predicts R^e, $R^e - R^f$, and σ_q.

6. Concluding Remarks

Our purpose in this paper is to provide researchers working with more complex model economies than the one studied here, with some guidance to help select from among the many available solution algorithms. We expect that in these more complex problems, computational speed and programming convenience will be important, desirable characteristics, in addition to accuracy. With this in mind, we compared and evaluated six computational algorithms for solving models with occasionally binding inequality constraints. These algorithms include: three versions of Marcet's parameterized expectations algorithm (PEA); a version of Judd's Spectral-Galerkin algorithm, extended here to include a Lagrange multiplier function as one of the objects sought; two finite element methods, Coleman's FEM-collocation algorithm, modified to accommodate a Lagrange multiplier, and McGrattan's FEM-Galerkin algorithm, which accommodates inequality constraints by including a penalty function in the

objective. In addition, to provide a benchmark solution, we also did dynamic programming on a discretized version of our model with a very fine grid. A unique feature of our analysis is that we illustrate the use of the Euler residual function in evaluating the accuracy of a solution algorithm.[46]

To our initial surprise, all the algorithms worked quite well. We were particularly surprised at the accuracy with which several of the algorithms predict the Lagrange multiplier. Even algorithms such as McGrattan's and a version of Marcet's, which compute the multiplier indirectly, provide reasonable estimates of this function. Also, for the most part all of the algorithms are reasonably accurate for computing particular statistics involving endogenous variables from the example model economy.

Still, we have developed information we believe is useful for discriminating among these algorithms. By far the easiest algorithm to implement is Marcet's PEA. As Marcet (1988) points out, the algorithm requires virtually no modification to accommodate inequality constraints. In the case of the other algorithms, accommodating inequality constraints involves substantial complications. For example, implementation of McGrattan's method requires considerable 'baby sitting' of the computer program, as one tries out various values of a penalty function parameter. The Spectral-Galerkin and FEM-collocation methods also entail additional complications to accommodate inequality constraints. This reflects the fact that they require directly parameterizing a Lagrange multiplier function, in addition to the policy functions.

While Marcet's PEA seems to be the easiest to implement, we had difficulties with conventional versions of it. A key component of those versions is a cumbersome nonlinear regression step, potentially involving tens of thousands of observations. One reason for the large number of observations is that the explanatory variables are inefficiently concentrated in a narrow range. We devised an alternative (PEA-collocation), in which the regression step is linear, the explanatory variables are orthogonal, and the required number of observations

[46]For an alternative procedure for evaluating the accuracy of a solution algorithm, see den Haan and Marcet (1994).

in the regression is very small: no more than sixteen in our experiments. This method produced results as accurate as the best other method, and is orders of magnitude faster.

Although it is clear that PEA-collocation is the best solution method for our example, that does not guarantee that it will dominate in higher dimensional cases. Here, there are at least two considerations. First, do the linearity and orthogonality properties of PEA-collocation survive into multidimensional settings? In appendix 2, we define multidimensional PEA-collocation and show that these properties do indeed survive in general. The second consideration involves the mapping from a parameterized expectation function to policy and multiplier functions, which is at the heart of any PEA. In our example, this mapping is trivial, but in higher dimensions it involves solving nonlinear equations. In principle, there could be examples in which this is very costly in programmer and/or computer time, in which case perhaps an alternative method might dominate. Here, it should be born in mind that PEA's have been applied routinely in high dimensional models (see footnote 5).

References

[1] Aiyagari, S. Rao. 1993. Frictions in Asset Pricing and Macroeconomics. Federal Reserve Bank of Minneapolis Working Paper No. 518, September.

[2] Aiyagari, S. Rao and Mark Gertler. 1991. Asset Returns With Transactions Costs and Uninsured Individual Risk. *Journal of Monetary Economics*, 27: 311-331.

[3] Atkeson, Andy and Patrick Kehoe. 1993. Industry Evolution and Transition: The Role of Information Capital. Manuscript.

[4] Bizer, David S., and Kenneth L. Judd, 1989. Uncertainty and Taxation. *American Economic Review, Papers and Proceedings*, May.

[5] Boldrin, Michele, Lawrence J. Christiano, and Jonas D.M. Fisher. 1994. Habit Persistence and the Equity Premium in the Neoclassical Growth Model. Manuscript.

[6] Chari, V.V., Lawrence J. Christiano, and Patrick Kehoe. Forthcoming. "Policy Analysis in Business Cycle Models", in Cooley, editor, *Frontiers of Business Cycle Research*, Princeton University Press, Princeton, NJ.

[7] Christiano, Lawrence J. 1991. Modeling the Liquidity Effect of a Monetary Shock. Federal Reserve Bank of Minneapolis *Quarterly Review*, Winter.

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

[9] Coleman, Wilbur John. 1988. Money, Interest and Capital in a Cash-in-Advance Economy. International Finance Discussion Paper No. 323. Washington: Board of Governors, Federal Reserve System.

[10] Coleman, Wilbur John, Christian Gilles, and Pamela Labadie. 1992. The Liquidity Premium in Average Interest Rates. *Journal of Monetary Economics*, 30: 449-465.

[11] Danthine, Jean-Pierre and John B. Donaldson. 1981. Stochastic Properties of Fast vs. Slow Growing Economies. *Econometrica* 49: 1007-1033.

[12] den Haan, Wouter. 1993. Heterogeneity, Aggregate Uncertainty and the Short Term Interest Rate: A Case Study of Two Solution Techniques. University of California at San Diego Manuscript.

[13] den Haan, Wouter and Albert Marcet. 1990. Solving the Stochastic Growth Model by Parameterizing Expectations. *Journal of Business and Economic Statistics* 8: 31-34.

[14] den Haan, Wouter and Albert Marcet. 1994. Accuracy in Simulation. *Review of Economic Studies.*

[15] Heaton, John and Deborah 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.

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

[17] Judd, Ken. 1991. Minimum Weighted Residuals Methods for Solving Aggregate Growth Models. Institute for Empirical Macroeconomics Discussion Paper No. 49.

[18] Judd, Ken. 1992a. Projection Methods for Solving Aggregate Growth Models. *Journal of Economic Theory*, 58: 410-452.

[19] Judd, Ken. 1992b. *Numerical Methods in Economics.* Manuscript.

[20] Judd, Ken. 1993. Comments on Marcet, Rust and Pakes. Forthcoming comments made at the World Congress of the Econometric Society, Barcelona.

[21] Judd, Ken and Antonio Bernardo. 1994. Projection Methods for Computing Conditional Expectations. Stanford University Working Paper.

[22] Kahn, James. 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.

[23] Kiyotaki, Nobuhiro and John Moore. 1993. Credit Cycles. Manuscript.

[24] Luenberger, David. 1969. *Optimization by Vector Space Methods.* New York: John Wiley & Sons.

[25] Marcet, Albert. 1988. Solving Nonlinear Stochastic Growth Models by Parametrizing Expectations. Carnegie-Mellon University, mimeo.

[26] Marcet, Albert and Juan Ketterer. 1989. Introducing Derivative Securities: A General Equilibrium Approach. Carnegie-Mellon University Working Paper.

[27] Marcet, Albert and Ramone Marimon. 1992. Communication, Committment and Growth. *Journal of Economic Theory*, 58: 219-249.

[28] Marcet, Albert and David Marshall. 1994. Solving Non-linear Rational Expectations Models by Parameterizing Expectations: Convergence to Stationary Solutions. Northwestern University Working Paper.

[29] Marcet, Albert and Ken Singleton. 1990. Equilibrium Asset Prices and Savings of Heterogeneous Agents In the Presence of Portfolio Constraints. Unpublished manuscript.

[30] Marshall, David. 1992. Inflation and Asset Returns in a Monetary Economy. *Journal of Finance*, 47: 1315-42.

[31] McCurdy, Tom and Nicholas Ricketts. 1992. An international economy with country-specific money and productivity growth processes. Queen's University Manuscript.

[32] McGrattan, Ellen. 1993. Solving the Stochastic Growth Model with a Finite Element Method. Federal Reserve Bank of Minneapolis mimeograph.

[33] Press, William H., Saul Teukolsky, William Vetterling, and Brian Flannery. 1992. *Numerical Recipes in Fortran.* New York: University of Cambridge Press.

[34] Reddy, J.N. 1993. *An Introduction to the Finite Element Method.* New York: McGraw-Hill.

[35] Sargent, Thomas. 1980. Tobin's q and the Rate of Investment in General Equilibrium. Carnegie-Rochester Conference Series on Public Policy 12.

[36] Stokey, Nancy and Robert E. Lucas, Jr., with Edward Prescott. 1989. *Recursive Methods in Economic Dynamics.* Cambridge, MA: Harvard University Press.

[37] Stroud, A.H. 1971. *Approximate Calculation of Multiple Integrals.* Englewood Cliffs, NJ: Prentice-Hall.

[38] Telmer, Chris. Forthcoming. Asset Pricing Puzzles and Incomplete Markets. *Journal of Finance.*

Appendix 1: The Dynamic Programming Algorithm

Our DP algorithm is standard. 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:

$$v_{j+1}(k,\theta) = \max_{k' \in A(k,\theta)} \left\{ u(k,k',\theta) + \frac{1}{2}\beta\left[v_j(k',\sigma) + v_j(k',-\sigma)\right] \right\},$$

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

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

and

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

for the constrained problem, and

$$A(k,\theta) = \vec{k} \cap \{k' : 0 \leq k' \leq \exp(\theta)k^{\alpha} + (1-\delta)k\}$$

for the unconstrained problem. 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 $k_i < k_{i+1}$, $i = 1,2,\ldots,M-1$, $k_1 = 16.9$, $k_M = 55.1$, and $M = 20{,}000$. 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.

$$G(k,\theta) = \underset{k' \in A(k,\theta)}{\operatorname{argmax}} \left\{ u(k,k',\theta) + \frac{1}{2}\beta\left[v(k',\sigma) + v(k',-\sigma)\right] \right\},$$

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

The DP investment decision rules graphed in section 5 and the DP second moment properties are based on $G(k,\theta)$. The DP version of the multiplier reported in section 5 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,

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

Appendix 2: Multidimensional Applications of PEA-Collocation

In this appendix we describe how PEA-collocation is applied in models with an arbitrary finite number of endogenous and exogenous state variables. We show that the principle qualitative features of PEA-collocation (e.g., linearity of the regression and orthogonality of the regressors) survive in a multidimensional setting. We also show that PEA-collocation encounters a 'curse of dimensionality' problem in very high dimensional systems. We propose an alternative, PEA-Galerkin, for dealing with circumstances like this.

The Problem

Let $k \in K \subset \Re^l$ denote a vector of endogenous state variables. We suppose that the exogenous variables, $\theta \in \Theta \subset \Re^m$, are a first order, stationary Markov process with transition density $p_{\theta'}(\theta' \mid \theta)$. Since we restrict m only to be finite, this is equivalent to assuming the exogenous variables have an arbitrary finite ordered Markov representation. In contrast to the analysis in the main text, here we suppose θ is a vector of continuous random variables. Let $u : K \times K \times \Theta \to \Re$ denote the one-period return function. This may be an indirect utility function that results after static decisions, such as labor supply in standard business cycle models or the sectoral allocation of capital in a multisector model, have been maximized out. We consider model economies which lead to the analysis of the following functional equation:[47]

$$W(k, \theta) = \max_{k' \in \Gamma(k,\theta), G(k,k',\theta) \geq 0_w} u(k, k', \theta) + \beta \int_{\theta' \in \Theta} W(k', \theta') p_{\theta'}(\theta' \mid \theta) d\theta', \text{ for all } k \in K, \theta \in \Theta$$

where G is a $w \times 1$ vector-valued function, $G : K \times K \times \Theta \to \Re^w$. Also, $W : K \times \Theta \to \Re$ is a value function, and $\Gamma : K \times \Theta \to K$ and $G(k, k', \theta) \geq 0_w$ characterize the constraints on the choice of k' in the maximization. Here, 0_w denotes a $w \times 1$ vector of zeros. The correspondence Γ summarizes constraints that either bind always or never, while the function G summarizes restrictions that bind occasionally.[48] In what follows it is convenient to use the notation $s = \text{vec}(k, \theta)$, where s is a $q \times 1$ vector, $q = l + m$.

Let $g : K \times \Theta \to K$ be the (single-valued) policy function which attains the maximum in the functional equation, and let $h : K \times \Theta \to \Re^w$ denote the $w \times 1$ vector-valued multiplier

[47] For a detailed discussion of model economies like this, see Stokey and Lucas with Prescott (1989).

[48] By a constraint never binding, we mean that its multiplier is zero for almost all $(k, \theta) \in K \times \Theta$. An example of this is the one sector growth model where Γ summarizes the nonnegativity constraints on consumption and capital, and the appropriate Inada conditions on u are satisfied. By a constraint always binding, we mean that its multiplier is almost always nonzero. An example of this is the resource constraint in the model just described. In the model analyzed in the body of the paper, Γ summarizes the resource constraint and nonnegativity constraints on capital and consumption, and the constraint, $G(k, k', \theta) \geq 0_w$, is the nonnegativity constraint on gross investment.

function corresponding to the constraints, $G(k,k',\theta) \geq 0_w$. We suppose that the policy and multiplier functions must satisfy the Euler equations,

$$u_2(k,g(k,\theta),\theta) + G_2(k,g(k,\theta),\theta)^T h(k,\theta) + \beta \int_{\theta'} m(k,\theta,\theta';g,h) p_{\theta'}(\theta' \mid \theta) d\theta' = 0_l, \quad (6.1)$$

and the Kuhn-Tucker conditions,

$$G(k,g(k,\theta),\theta) \geq 0_w, h(k,\theta) \geq 0_w, \text{and } h(k,\theta) .* G(k,g(k,\theta),\theta) = 0_w, \quad (6.2)$$

where $.*$ denotes element-by-element multiplication. Also,

$$m(k,\theta,\theta';g,h) = u_1(g(k,\theta), g(g(k,\theta),\theta'),\theta') + G_1(g(k,\theta), g(g(k,\theta),\theta'),\theta')^T h(g(k,\theta),\theta'), \quad (6.3)$$

where $m : K \times \Theta \times \Theta \to \Re^l$, is an $l \times 1$ vector-valued function. In (6.1)-(6.3), u_i denotes the $l \times 1$ vector of derivatives of u with respect to the i^{th} argument, G_i denotes the $w \times l$ matrix of derivatives of G with respect to its i^{th} argument, and ^{T} denotes the matrix transposition operator. The problem is to approximate g and h, solutions to the functional equations (6.1)-(6.3).

The PEA

The PEA approximates g and h indirectly by parameterizing the j-th conditional expectation in (6.1) by a function $\exp(\hat{e}_{a^j}(s))$, $j = 1,2,\ldots,l$:

$$\exp(\hat{e}_{a^j}(s)) \approx \int_{\theta' \in \Theta} m^j(s,\theta';g,h) p_{\theta'}(\theta' \mid \theta) d\theta', j = 1,2,\ldots,l, \quad (6.4)$$

for all $s \in K \times \Theta$. Here, $a^j \in \Re^{Np}$ is a finite set of parameters, and m^j is the j^{th} element of the function, m. Let $a = \text{vec}(a^1, a^2, \ldots, a^l)$. We define a mapping from a to policy and multiplier functions, $\hat{g}_a$ and $\hat{h}_a$, as follows. For any given a, replace the conditional expectation in (6.1) by the parameterized expectation in (6.4) and let $\hat{g}_a$ and $\hat{h}_a$ denote the policy and multiplier functions which satisfy the Euler equation and Kuhn-Tucker conditions.[49] These policy and multiplier functions then imply a conditional expectation function, $\int_{\theta' \in \Theta} m(s,\theta';\hat{g}_a,\hat{h}_a) p_{\theta'}(\theta' \mid \theta) d\theta'$ for all $s \in K \times \Theta$. A new vector of parameters, $a' = S(a; l \cdot N^p)$, is then chosen to make the function $\exp(\hat{e}_{a'}(s))$ as close as possible to this conditional expectation. Here, lN^p denotes the number of elements in a. A PEA seeks an a^* such that $a^* = S(a^*; l \cdot N^p)$.

PEA-Collocation

As in the text, we construct $\hat{e}_a(s)$ using Chebyshev polynomials. We begin by defining what these are in a multidimensional setting, and by displaying their discrete orthogonality property. As in the text, it is this property which accounts for the orthogonal regressor property of PEA-collocation.

[49] In general, finding $\hat{g}_a$ and $\hat{h}_a$ for any $(k,\theta) \in K \times \Theta$ requires solving a system of equations using numerical methods. In contrast, in the model economy of the main text, the solution to these equations has a simple analytic form (see (4.8) and (4.9)).

Chebyshev polynomials with q-dimensional domain are constructed from the one dimensional Chebyshev polynomials studied in the text. Let $\Phi_n^{(1)} = \{T_0(x), T_1(x), \ldots, T_{n-1}(x)\}$ denote the basis functions for one dimensional Chebyshev polynomials of degree $n-1$, for $n > 1$, where $x \in (-1,1)$. The *tensor product basis* for degree $n-1$ Chebyshev polynomial functions of q variables is constructed by taking all possible q-term products of the n elements in $\Phi_n^{(1)}$. Accordingly, the resulting basis is

$$\Phi_n^{(q)} = \left\{ T_{i_1}(x_1) T_{i_2}(x_2) \cdots T_{i_q}(x_q) \mid i_j = 0, 1, \ldots, n-1, j = 1, 2, \ldots, q \right\}.$$

Here, $(x_1, x_2, ..., x_q)$ is an element of the q-fold Cartesian product of $(-1,1)$, which we denote by $(-1,1)^q$. Notice that $\Phi_n^{(q)}$ contains n^q elements.

A convenient feature of this tensor product basis is that it inherits the discrete orthogonality properties of $\Phi_n^{(1)}$ (see Judd (1992b, chapter 5) and the references given there). Let $\phi_1, \phi_2, \ldots, \phi_{n^q}$ be a list of the elements of $\Phi_n^{(q)}$, where $\phi_v : (-1,1)^q \to (-1,1)$ for $v = 1, ..., n^q$. Then the discrete orthogonality property in the multidimensional setting is, for $i, j < n^q$,

$$\sum_{v=1}^{n^q} \phi_i(\vec{r}_v)\phi_j(\vec{r}_v) = \begin{cases} 0, \text{ for } i \neq j \\ c_i(n,q), \text{ for } i = j \end{cases},$$

where $c_i(n,q)$ are constants that depend on the basis and $\vec{r}_v \in (-1,1)^q$ is composed of a selection of q elements from the set of n zeros of T_n, $v = 1, 2, \ldots, n^q$. In particular, the set of $\vec{r}_v$'s is defined by the n^q ways of choosing q of the n zeros of T_n. The zeros of T_n are given by

$$r_k = \cos\left(\frac{(2k-1)\pi}{2n}\right), k = 1, \ldots, n.$$

Also, for $i = 1, 2, \ldots, n^q$

$$c_i(n,q) = \sum_{v=1}^{n^q} \phi_i(\vec{r}_v)^2.$$

We construct the parameterized expectation function using the elements of $\Phi_n^{(q)}$ as follows:

$$\widehat{\epsilon}_{a^j}(s) = \sum_{i=1}^{n^q} a_i^j \phi_i(\varphi(s)), j = 1, 2, \ldots, l.$$

Notice that $N^p = n^q$ here. The function φ is the multidimensional version of the analogous function used in the main text (see (3.6).) That is $\varphi : \prod_{i=1}^q (\underline{s}_i, \overline{s}_i) \subset \Re^q \to (-1,1)^q$, where $(\underline{s}_i, \overline{s}_i), i = 1, 2, \ldots, q$ bound the exogenous and endogenous variables.

We now derive the multidimensional version of the orthogonal regressor result, (3.15). With PEA-collocation, $a' = S(a; l \cdot n^q)$ is defined by:

$$R_j(s_v; \widehat{g}_{a'}, \widehat{h}_{a'}) = \exp(\widehat{\epsilon}_{a'^j}(\varphi(s_v))) - \int_{\theta' \in \Theta} m^j(s_v, \theta'; \widehat{g}_a, \widehat{h}_a) p_{\theta'}(\theta' \mid \theta) d\theta' = 0, j = 1, 2, \ldots, l \tag{6.5}$$

for $v = 1, \ldots, n^q$, where $s_v = \varphi^{-1}(\vec{r}_v)$, $v = 1, \ldots, n^q$.[50] Now (6.5) holds if and only if it holds for the log of the terms on each side of the minus sign. That is, for each $v = 1, 2, \ldots, n^q$ and each $j = 1, 2, \ldots, l$, (6.5) holds if and only if

$$\sum_{i=1}^{n^q} a_i^{'j} \phi_i(\varphi(s_v)) = \ln \left(\int_{\theta' \in \Theta} m^j(s_v, \theta'; \hat{g}_a, \hat{h}_a) p_{\theta'}(\theta' \mid \theta) d\theta' \right). \tag{6.6}$$

Multiply both sides of each equation in (6.6) by $\phi_1(s_v)$ and for fixed j sum over $v = 1, 2, \ldots, n^q$. By the discrete orthogonality property, all terms on the left side of the equality, except those involving $\phi_1(\varphi(s_v))^2$, $v = 1, \ldots, n^q$, are zero. Repeating this procedure for $\phi_2, \phi_3, \ldots, \phi_{n^q}$, one finds that, analogous to (3.15), the mapping $a' = S(a; l \cdot n^q)$ has the simple analytical form,

$$a_i^{'j} = \frac{1}{c_i(n,q)} \sum_{v=1}^{n^q} \phi_i(\varphi(s_v)) \ln \left(\int_{\theta' \in \Theta} m^j(s_v, \theta'; \hat{g}_a, \hat{h}_a) p_{\theta'}(\theta' \mid \theta) d\theta' \right), i = 1, \ldots, n^q \tag{6.7}$$

for $j = 1, 2, \ldots, l$.

PEA-Galerkin

A disadvantage of a tensor product basis is that the number of elements in the basis grows exponentially as the dimension increases. One could instead work with a strict subset of the number of elements in the tensor product basis. For example, Judd (1992b, chapter 5) suggests working with the following subset:

$$C_n^{(q)} = \left\{ T_{i_1}(x_1) T_{i_2}(x_2) \cdots T_{i_q}(x_q) \mid \sum_{j=1}^{q} i_j \leq n - 1, \; i_1, i_2, \ldots, i_q \geq 0 \right\}.$$

Notice that $C_n^{(q)} \subset \Phi_n^{(q)}$, since $C_n^{(q)}$ simply deletes high-order cross product terms in $\Phi_n^{(q)}$. For very large problems the computational burden of finding the a^* that solves $a^* = S(a^*; l \cdot n^q)$ may be unacceptable. In these circumstances, a useful alternative may be to use a smaller basis. However, if one continues to work with the Chebyshev zeros, $\vec{r}_v, v = 1, 2, \ldots, n^q$, then PEA-collocation is no longer implementable in general. This is because PEA-collocation now attempts to solve the n^q equations, (6.6), using less than n^q unknowns. There are several options. One is to apply PEA-collocation to a reduced number of equations. Another is to maintain the number of equations and apply a different weighted residual method.

From the text, it is clear that there are many such methods. One such method is a modified version of the Galerkin procedure discussed in the text, *PEA-Galerkin*, which is applied as follows. First, select a value for n, and choose a subset of N^p basis functions from $\Phi_n^{(q)}$, where $N^p < n^q$. Then, compute $\vec{r}_v, v = 1, 2, \ldots, n^q$ as above and form the $N^p \times n^q$

[50] The integral in (6.5) could be approximated by quadrature or Monte Carlo methods.

matrix A and the $n^q \times 1$ vector $\vec{R}_j(\vec{s}, \hat{g}_a, \hat{h}_a)$ as follows:

$$A = \begin{bmatrix} \phi_1(\vec{r}_1) & \phi_1(\vec{r}_2) & \cdots & \phi_1(\vec{r}_{n^q}) \\ \phi_2(\vec{r}_1) & \phi_2(\vec{r}_2) & \cdots & \phi_2(\vec{r}_{n^q}) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_{N^p}(\vec{r}_1) & \phi_{N^p}(\vec{r}_2) & \cdots & \phi_{N^p}(\vec{r}_{n^q}) \end{bmatrix}, \vec{R}_j(\vec{s}, \hat{g}_a, \hat{h}_a) = \begin{bmatrix} R_j(s_1, \hat{g}_a, \hat{h}_a) \\ R_j(s_2, \hat{g}_a, \hat{h}_a) \\ \vdots \\ R_j(s_{n^q}, \hat{g}_a, \hat{h}_a) \end{bmatrix},$$

where $\vec{s} = [s_1, \ldots, s_{n^q}]$, and s_v, $v = 1, 2, \ldots, n^q$, is as defined above. By the discrete orthogonality property, the rows of A are orthogonal. Finally, find the value of a that solves the system of $l \cdot N^p$ nonlinear equations:

$$A\vec{R}_j(\vec{s}, \hat{g}_a, \hat{h}_a) = 0, j = 1, 2, \ldots, l.$$

Table 1

Computation Times and Maximum Absolute Values
of the Euler Residuals for the Various Algorithms

Model	Approximation Method		Time (Seconds)	MAV of Euler Residuals $\theta = \sigma$	MAV of Euler Residuals $\theta = -\sigma$
Unconstrained	PEA:	$N^p = 3$	155.1	9.2×10^{-5}	1.3×10^{-4}
				(1.0×10^{-4})	(1.6×10^{-3})
		$N^p = 3^*$	151.9	3.3×10^{-5}	1.6×10^{-4}
				(5.1×10^{-4})	(1.5×10^{-3})
		$N^p = 3^{**}$	0.6	1.7×10^{-6}	2.1×10^{-6}
				(3.2×10^{-6})	(4.1×10^{-6})
	Spectral-Galerkin:	$N^J = 3$	3.5	4.5×10^{-4}	1.2×10^{-4}
		$N^J = 5$	6.3	3.7×10^{-7}	6.2×10^{-7}
	FEM-Collocation:	$N^f = 36$	253.1	1.0×10^{-4}	7.8×10^{-5}
		$N^f = 72$	557.6	3.3×10^{-5}	2.7×10^{-5}
	FEM-Galerkin:	$N^f = 18$	8.8	3.7×10^{-4}	1.3×10^{-4}
		$N^f = 36$	17.9	7.2×10^{-5}	4.8×10^{-5}
Constrained	PEA:	$N^p = 3$	174.4	1.1×10^{-4}	1.7×10^{-4}
				(1.6×10^{-4})	(8.8×10^{-4})
		$N^p = 3^*$	181.7	3.3×10^{-5}	1.9×10^{-4}
				(5.6×10^{-4})	(7.2×10^{-3})
		$N^p = 8^{**}$	2.5	2.6×10^{-5}	9.9×10^{-5}
				(2.6×10^{-5})	(9.9×10^{-5})
	Spectral-Galerkin:	$(5,3,2)$	17.6	6.2×10^{-4}	7.3×10^{-4}
		$(9,5,4)$	27.9	8.8×10^{-5}	8.0×10^{-5}
	FEM-Collocation:	$N^f = 36$	522.9	1.0×10^{-4}	7.8×10^{-5}
		$N^f = 72$	996.2	3.4×10^{-5}	2.8×10^{-5}
	FEM-Galerkin:	$N^f = 18$	212.5	7.3×10^{-3}	2.4×10^{-2}
		$N^f = 36$	1783.9	1.5×10^{-4}	3.7×10^{-4}

Notes: (i) For the Spectral-Galerkin and FEM-Galerkin approximations to the constrained model we used the unconstrained approximations for starting values. Starting values for the unconstrained calculations were based on the log-linear approximate decision rules. The times displayed include computation times for these starting values. (ii) In practice, the FEM-Galerkin algorithm as applied to the constrained model involves solving penalty function versions of the model for several values of the penalty parameter π. The computation times for this application reflect this fact. (iii) The ordered triplets for the Spectral-Galerkin approximations to the constrained model correspond to $(N(\sigma), N(-\sigma), N^\lambda)$. (iv) See the text for an explanation of the asterisks associated with the PEA entries. (v) The MAV numbers corresponding to the PEA have the following interpretation. They are based on confidence intervals from simulations of length 10,000 based on implied decision rules for the approximation in question. The numbers not in parenthesis correspond to MAVs based on an interval containing 90% of the realizations of the capital stock. The numbers in parenthesis correspond to MAVs based on an interval that contains 100% of the realizations.

Table 2

Statistics from Various Approximations of the Unconstrained Model

		PEA: $N^p =$			S-G: $N^J =$		FEM-C: $N^f =$		FEM-G: $N^f =$	
Statistic	DP	3	3*	3**	3	5	36	72	18	36
R^e	3.099	3.120	3.092	3.090	3.146	3.090	3.073	3.088	3.102	3.092
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
R^f	3.052	3.031	3.040	3.043	3.022	3.043	3.049	3.043	3.039	3.042
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
$R^e - R^f$	0.047	0.089	0.051	0.047	0.012	0.047	0.024	0.045	0.063	0.049
	(0.010)	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
σ_y	62.2	62.2	62.2	62.2	62.1	62.2	62.3	62.2	62.2	62.2
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
σ_c	7.22	7.21	7.25	7.13	7.31	7.13	7.15	7.13	7.12	7.13
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
σ_i	60.1	59.9	60.0	60.1	59.8	60.1	60.3	60.1	60.0	60.1
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
$\rho(y, c)$	0.354	0.362	0.361	0.358	0.370	0.358	0.357	0.358	0.358	0.358
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\rho(y, i)$	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993	0.993
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Freq($i < 0$)	8.89	9.89	9.17	8.64	10.50	8.72	8.23	8.59	8.97	8.67
	(0.36)	(0.38)	(0.37)	(0.37)	(0.39)	(0.37)	(0.35)	(0.36)	(0.37)	(0.36)

Notes: (i) Statistics shown are averages from samples of length 114 replicated 500 times. (ii) Freq($i < 0$) indicates the per cent rate at which gross investment is negative across samples. (iii) Numbers in parenthesis are Monte Carlo standard errors. (iv) See the text for a description of the asterisk notation used for the PEA entries in the table. (v) Finally, S-G stands for Spectral-Galerkin, FEM-C stands for FEM-Collocation and FEM-G stands for FEM-Galerkin.

Table 3

Statistics from Various Approximations of the Constrained Model

		PEA: $N^p =$			Spectral-Galerkin		FEM-C: $N^f =$		FEM-G: $N^f =$	
Statistic	DP	3	3*	8**	(5,3,2)	(9,5,4)	36	72	18	36
R^e	3.126	3.160	3.126	3.128	3.121	3.125	3.109	3.122	3.588	3.125
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.006)	(0.017)	(0.006)
R^f	3.055	3.043	3.052	3.053	3.051	3.053	3.058	3.054	3.069	3.053
	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.027)	(0.017)
$R^e - R^f$	0.071	0.117	0.074	0.076	0.070	0.072	0.051	0.068	0.519	0.072
	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.013)	(0.015)	(0.043)	(0.015)
σ_y	62.2	62.2	62.2	62.2	62.2	62.2	62.3	62.2	61.9	62.2
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
σ_c	7.08	7.15	7.20	7.09	7.09	7.08	7.10	7.07	8.96	7.07
	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.11)	(0.08)
σ_i	59.7	59.6	59.7	59.7	59.7	59.8	59.8	59.8	58.3	59.8
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.03)
σ_q	0.318	0.341	0.319	0.312	0.283	0.310	0.295	0.302	1.070	0.303
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.040)	(0.014)
$\rho(y,c)$	0.400	0.411	0.404	0.400	0.409	0.399	0.396	0.398	0.445	0.398
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)
$\rho(y,i)$	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.994	0.990	0.994
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
$\rho(y,q)$	0.229	0.248	0.236	0.226	0.201	0.220	0.217	0.220	0.157	0.218
	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)	(0.006)	(0.007)
Freq($\lambda > 0$)	9.67	10.92	10.11	9.73	8.08	9.31	9.04	9.31	4.53	9.17
	(0.39)	(0.41)	(0.40)	(0.40)	(0.36)	(0.39)	(0.38)	(0.39)	(0.22)	(0.39)

Notes: (i) Statistics shown are averages from samples of length 114 replicated 500 times. (ii) Freq($\lambda > 0$) indicates the per cent rate at which the constraint binds across samples. (iii) Numbers in parenthesis are Monte Carlo standard errors. (iv) See the text for a description of the asterisk notation used for the PEA entries in the table. (v) FEM-C stands for FEM-Collocation and FEM-G stands for FEM-Galerkin. (vi) The ordered triplets is the Spectral-Galerkin columns correspond to $(N(\sigma), N(-\sigma), N^{\lambda})$.

FIGURE 1

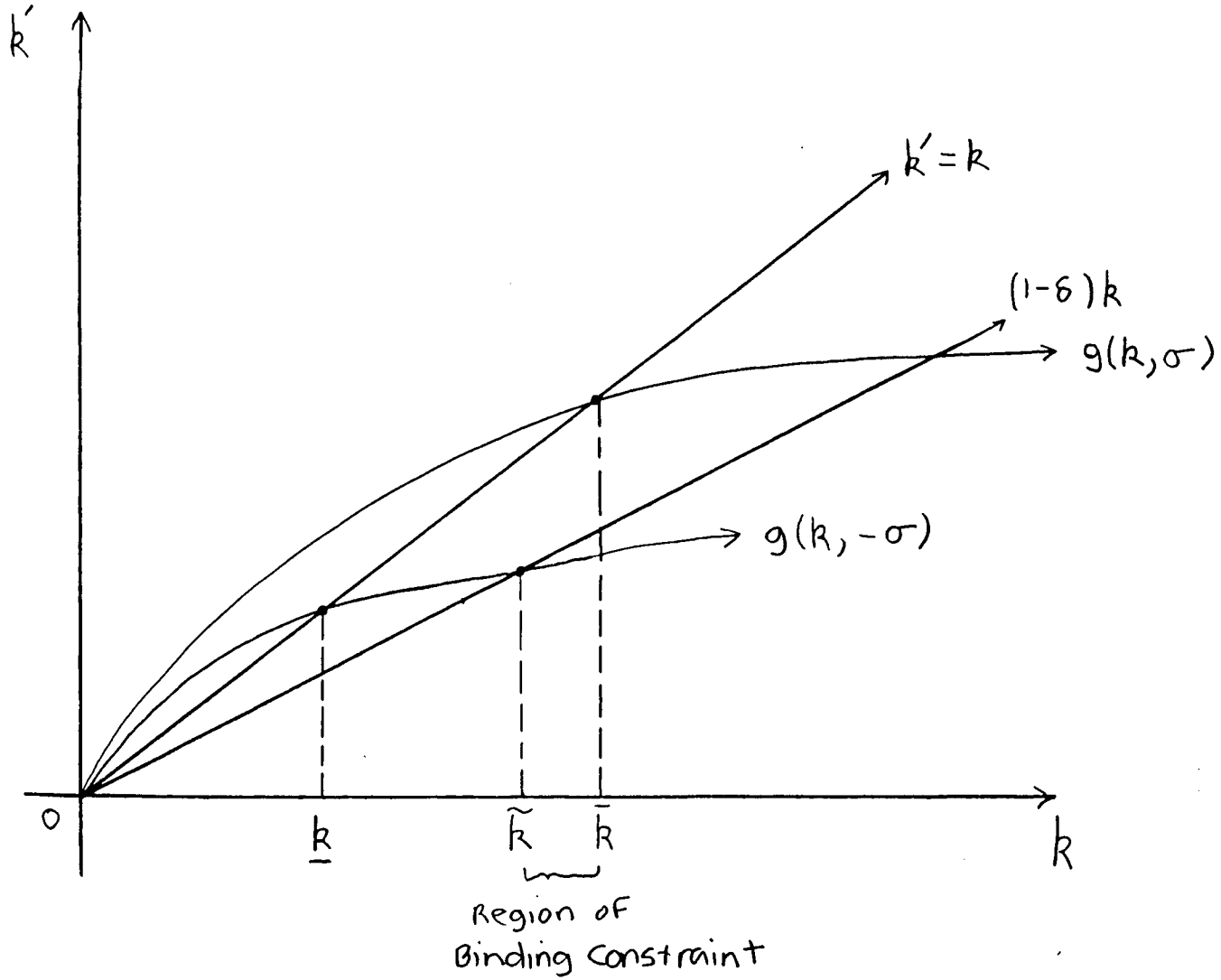


Figure 2a. Euler residuals for PEA and Spectral-Galerkin

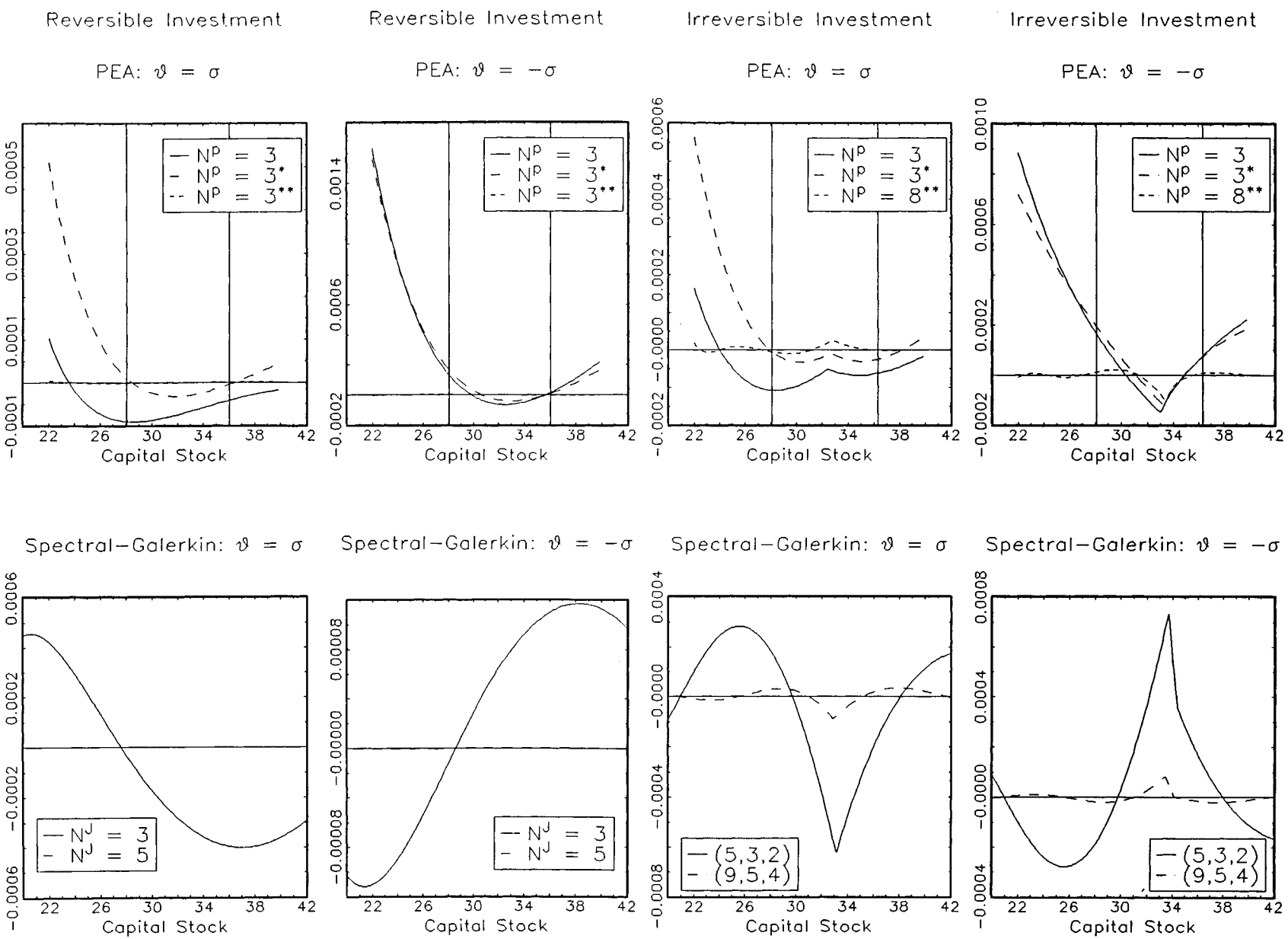


Figure 2b. Euler residuals for FEM-collocation and FEM-Galerkin

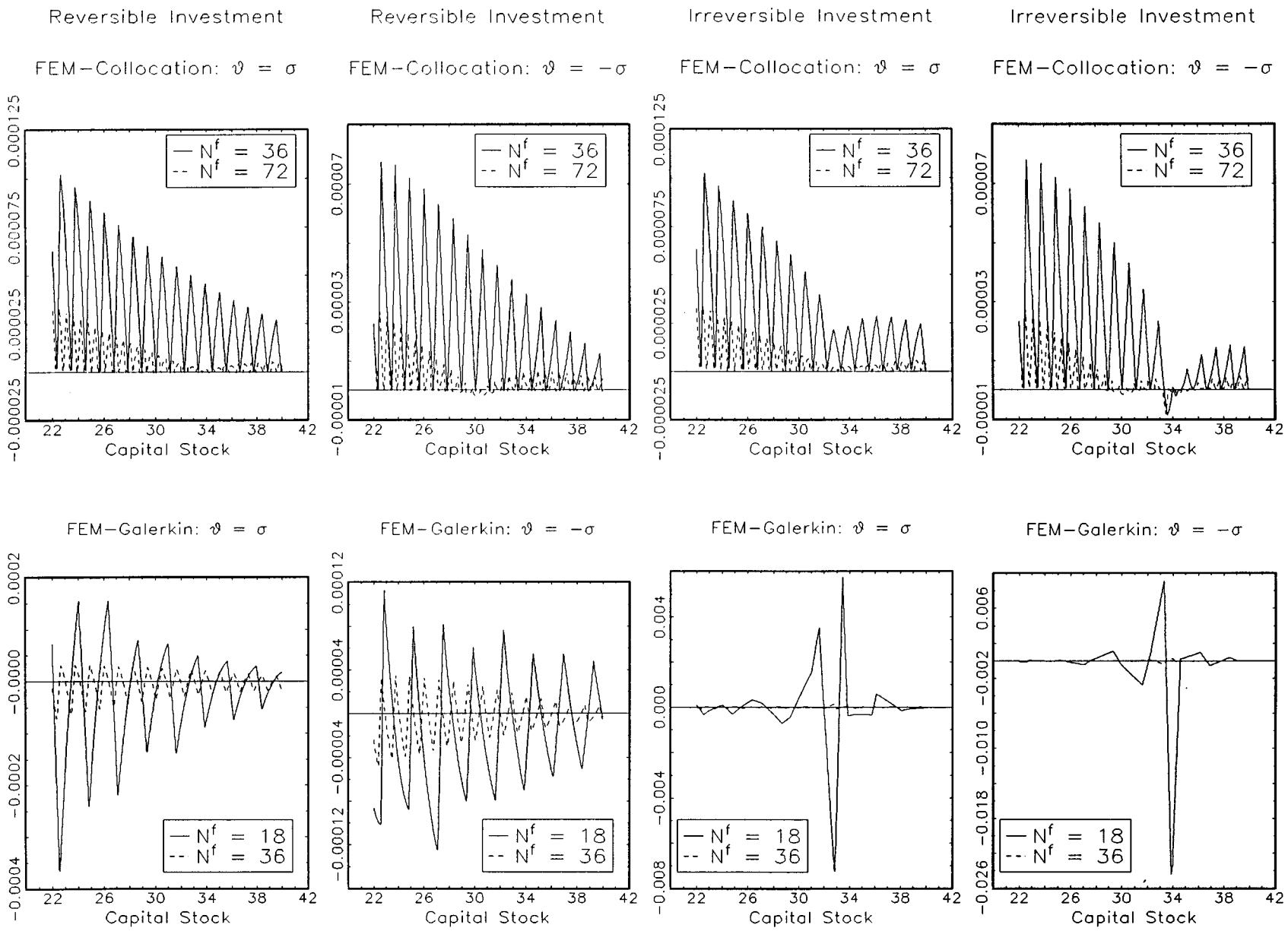


Figure 3. Policy functions implied by Dynamic Programming and Spectral-Galerkin

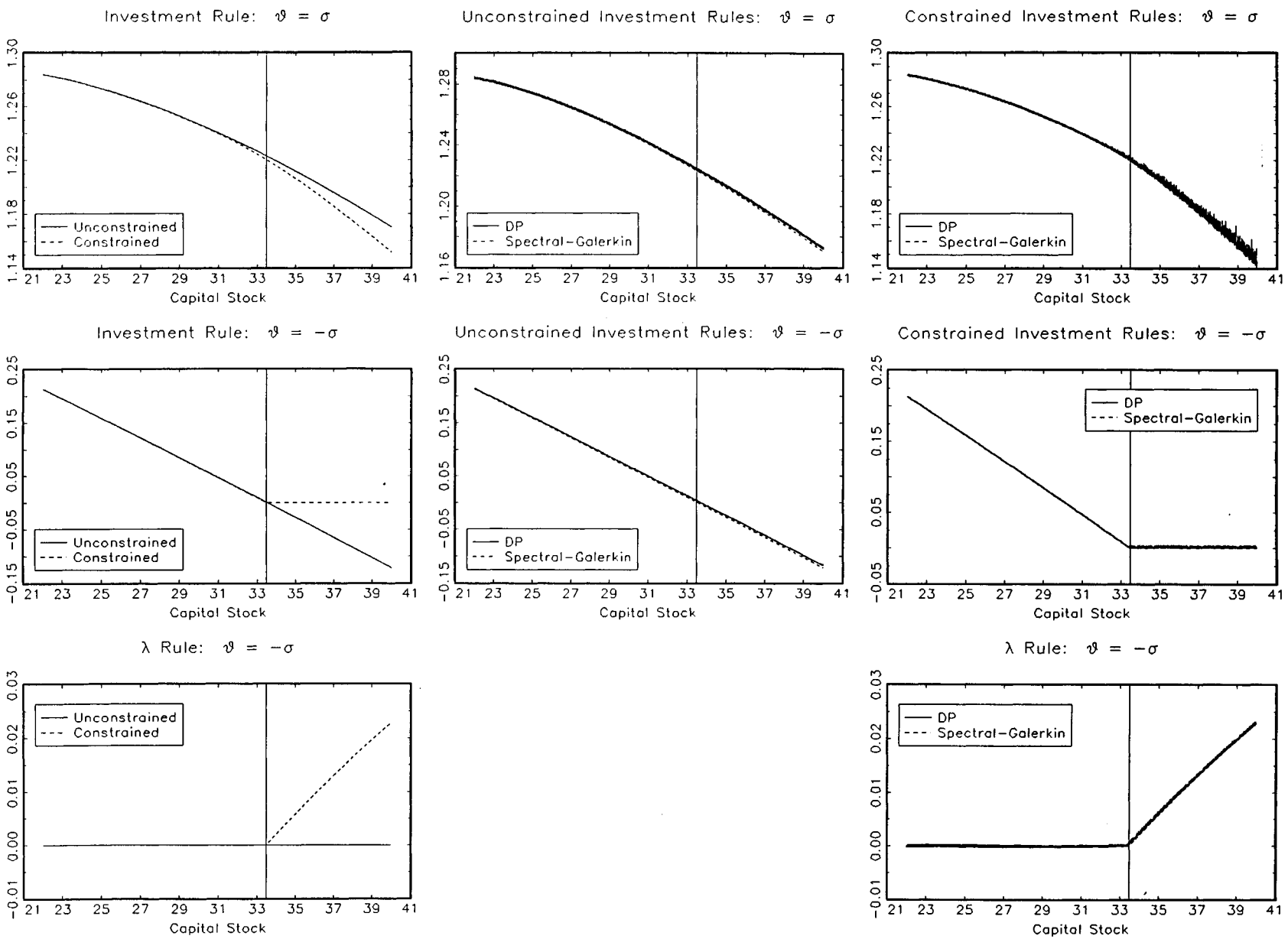


Figure 4. PEA versus Spectral-Galerkin Investment Rules in the Reversible Investment economy

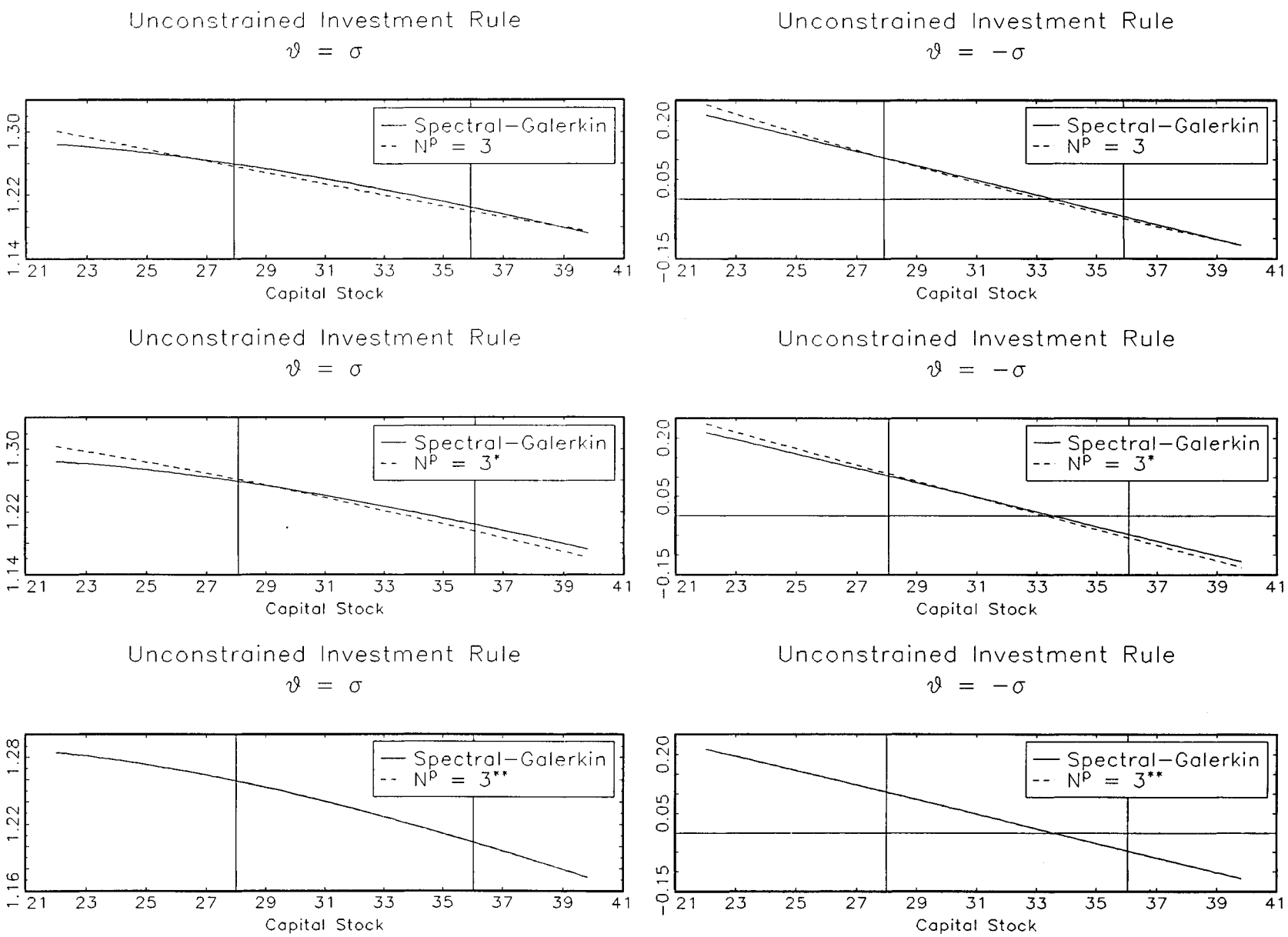
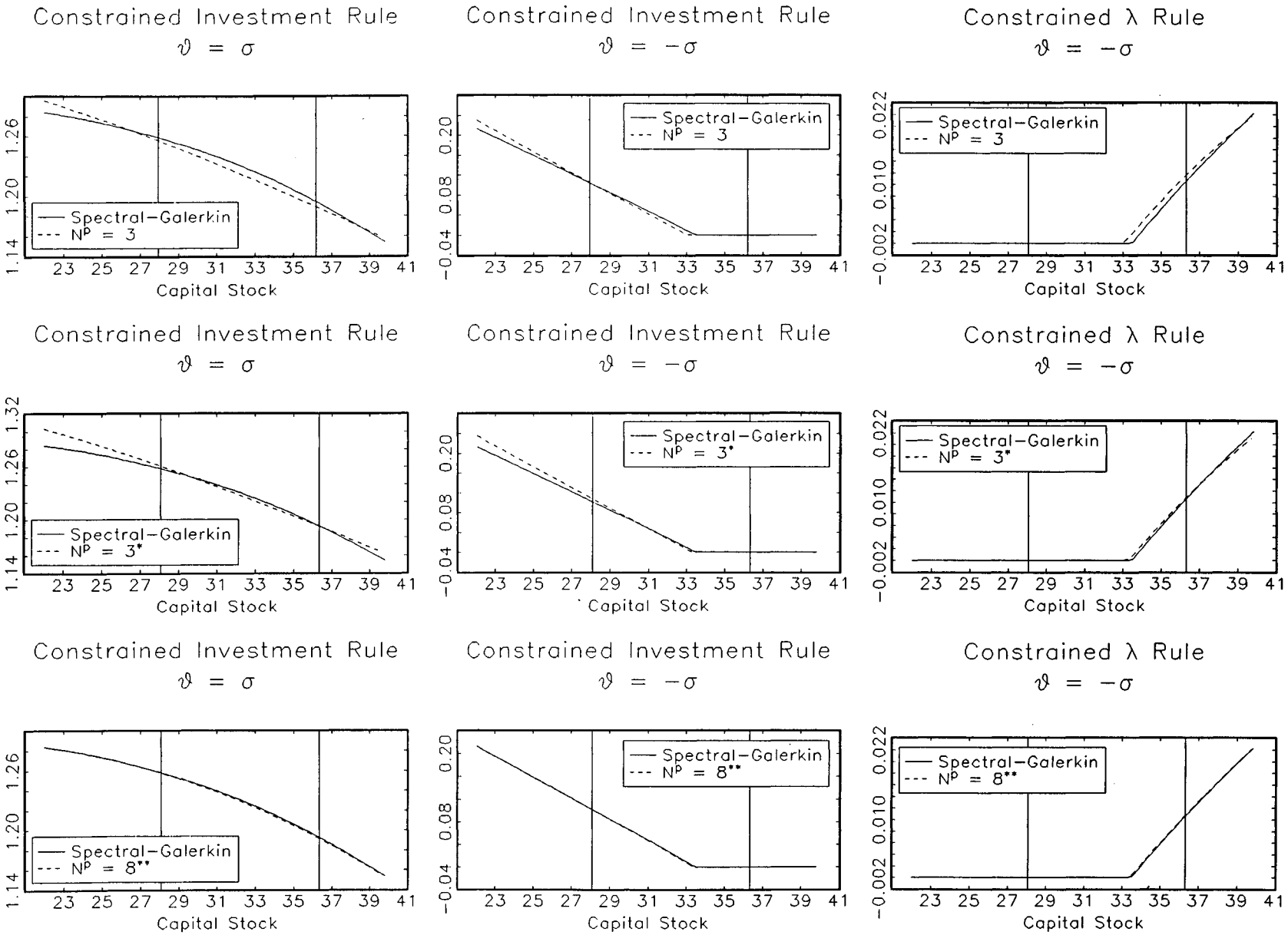


Figure 5. PEA versus Spectral-Galerkin Investment and Lambda Rules in the Irreversible Investment Economy



Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

REGIONAL ECONOMIC ISSUES

Estimating Monthly Regional Value Added by Combining Regional Input With National Production Data **WP-92-8**
Philip R. Israilevich and Kenneth N. Kuttner

Local Impact of Foreign Trade Zone **WP-92-9**
David D. Weiss

Trends and Prospects for Rural Manufacturing **WP-92-12**
William A. Testa

State and Local Government Spending--The Balance Between Investment and Consumption **WP-92-14**
Richard H. Mattoon

Forecasting with Regional Input-Output Tables **WP-92-20**
P.R. Israilevich, R. Mahidhara, and G.J.D. Hewings

A Primer on Global Auto Markets **WP-93-1**
Paul D. Ballew and Robert H. Schnorbus

Industry Approaches to Environmental Policy in the Great Lakes Region **WP-93-8**
David R. Allardice, Richard H. Mattoon and William A. Testa

The Midwest Stock Price Index--Leading Indicator of Regional Economic Activity **WP-93-9**
William A. Strauss

Lean Manufacturing and the Decision to Vertically Integrate Some Empirical Evidence From the U.S. Automobile Industry **WP-94-1**
Thomas H. Klier

Domestic Consumption Patterns and the Midwest Economy **WP-94-4**
Robert Schnorbus and Paul Ballew

ISSUES IN FINANCIAL REGULATION

Stock Margins and the Conditional Probability of Price Reversals **WP-93-5**
Paul Kofman and James T. Moser

Is There Lif(f)e After DTB?
Competitive Aspects of Cross Listed Futures
Contracts on Synchronous Markets **WP-93-11**
Paul Kofman, Tony Bouwman and James T. Moser

Opportunity Cost and Prudentiality: A Representative-
Agent Model of Futures Clearinghouse Behavior **WP-93-18**
Herbert L. Baer, Virginia G. France and James T. Moser

The Ownership Structure of Japanese Financial Institutions **WP-93-19**
Hesna Genay

Origins of the Modern Exchange Clearinghouse: A History of Early
Clearing and Settlement Methods at Futures Exchanges **WP-94-3**
James T. Moser

The Effect of Bank-Held Derivatives on Credit Accessibility **WP-94-5**
Elijah Brewer III, Bernadette A. Minton and James T. Moser

MACROECONOMIC ISSUES

An Examination of Change in Energy Dependence and Efficiency
in the Six Largest Energy Using Countries--1970-1988 **WP-92-2**
Jack L. Hervey

Does the Federal Reserve Affect Asset Prices? **WP-92-3**
Vefa Tarhan

Investment and Market Imperfections in the U.S. Manufacturing Sector **WP-92-4**
Paula R. Worthington

Business Cycle Durations and Postwar Stabilization of the U.S. Economy **WP-92-6**
Mark W. Watson

A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues
and Recent Performance **WP-92-7**
James H. Stock and Mark W. Watson

Working paper series continued

Production and Inventory Control at the General Motors Corporation During the 1920s and 1930s *Anil K. Kashyap and David W. Wilcox*	WP-92-10
Liquidity Effects, Monetary Policy and the Business Cycle *Lawrence J. Christiano and Martin Eichenbaum*	WP-92-15
Monetary Policy and External Finance: Interpreting the Behavior of Financial Flows and Interest Rate Spreads *Kenneth N. Kuttner*	WP-92-17
Testing Long Run Neutrality *Robert G. King and Mark W. Watson*	WP-92-18
A Policymaker's Guide to Indicators of Economic Activity *Charles Evans, Steven Strongin, and Francesca Eugeni*	WP-92-19
Barriers to Trade and Union Wage Dynamics *Ellen R. Rissman*	WP-92-22
Wage Growth and Sectoral Shifts: Phillips Curve Redux *Ellen R. Rissman*	WP-92-23
Excess Volatility and The Smoothing of Interest Rates: An Application Using Money Announcements *Steven Strongin*	WP-92-25
Market Structure, Technology and the Cyclicality of Output *Bruce Petersen and Steven Strongin*	WP-92-26
The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle *Steven Strongin*	WP-92-27
Earnings Losses and Displaced Workers *Louis S. Jacobson, Robert J. LaLonde, and Daniel G. Sullivan*	WP-92-28
Some Empirical Evidence of the Effects on Monetary Policy Shocks on Exchange Rates *Martin Eichenbaum and Charles Evans*	WP-92-32

Working paper series continued

An Unobserved-Components Model of Constant-Inflation Potential Output *Kenneth N. Kuttner*	WP-93-2
Investment, Cash Flow, and Sunk Costs *Paula R. Worthington*	WP-93-4
Lessons from the Japanese Main Bank System for Financial System Reform in Poland *Takeo Hoshi, Anil Kashyap, and Gary Loveman*	WP-93-6
Credit Conditions and the Cyclical Behavior of Inventories *Anil K. Kashyap, Owen A. Lamont and Jeremy C. Stein*	WP-93-7
Labor Productivity During the Great Depression *Michael D. Bordo and Charles L. Evans*	WP-93-10
Monetary Policy Shocks and Productivity Measures in the G-7 Countries *Charles L. Evans and Fernando Santos*	WP-93-12
Consumer Confidence and Economic Fluctuations *John G. Matsusaka and Argia M. Sbordone*	WP-93-13
Vector Autoregressions and Cointegration *Mark W. Watson*	WP-93-14
Testing for Cointegration When Some of the Cointegrating Vectors Are Known *Michael T. K. Horvath and Mark W. Watson*	WP-93-15
Technical Change, Diffusion, and Productivity *Jeffrey R. Campbell*	WP-93-16
Economic Activity and the Short-Term Credit Markets: An Analysis of Prices and Quantities *Benjamin M. Friedman and Kenneth N. Kuttner*	WP-93-17
Cyclical Productivity in a Model of Labor Hoarding *Argia M. Sbordone*	WP-93-20

Working paper series continued

The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds **WP-94-2**
Lawrence J. Christiano, Martin Eichenbaum and Charles Evans

Algorithms for Solving Dynamic Models with Occasionally Binding Constraints **WP-94-6**
Lawrence J. Christiano and Jonas D.M. Fisher