The Parametric Path Method for Solving Perfect Foresight Models

KENNETH L. JUDD
HOOVER INSTITUTION
STANFORD, CA 94305
JUDD@HOOVER.STANFORD.EDU

January, 1999

ABSTRACT. The parametric path method applies projection methods to compute the equilibrium time path of economic variables in large macroeconomic models. We exploit the special structure of economic time paths common in such models to reduce dimensionality. An illustrative example shows that the method can find excellent approximations with little computational cost.

Large-scale dynamic general equilibrium models are increasingly used in analyses of economic problems. However, their use is limited by numerical difficulty of solving such models. The perfect foresight aspect of dynamic general equilibrium analysis creates links between current and future economic variables. This simultaneity automatically generates large complex systems of nonlinear equations, and makes conventional computational general equilibrium procedures like Scarf's algorithm or homotopy procedures impractical.

We propose an algorithm which uses standard methods from numerical functional analysis and exploits the special structure of many dynamic general equilibrium models. While there are an infinite number of unknowns in an infinite-horizon general equilibrium model, the dynamic path is relatively well behaved. Specifically, dynamic general equilibrium analyses often assume convergence to a steady state, or, more generally, convergence to some known (or easily computed) dynamic path. During some initial phase, that convergence need not be well-behaved but asymptotically the convergence is governed by linear approximations about the asymptotic path. The idea of the parametric path method is to express the time path of economic variables as some function of time where the number of free parameters in the parameterization is far less than the number of unknown prices and quantities in the infinite-horizon economic model[1]. As long as the asymptotic behavior of equilibrium is well-behaved, it is possible to construct flexible and parsimonious parameterizations which can accurately approximate equilibrium. This reduction in the number

[1]It is common in numerical rational expectations models to parameterize the critical unknown functions; see Gustafson (1958), Wright and Williams (1982, 1984), and Miranda (1987) for the seminal contributions to this literature.

of unknowns substantially reduces the complexity of the numerical problem since the cost of solving methods for nonlinear equations is quadratic (at best) in the number of unknowns.

We first describe the general perfect foresight model and traditional solution methods. We then use the projection method (see Judd, 1992) to develop the parametric path method for solving perfect foresight models. We finish with a detailed application to a familiar simple perfect foresight model.

1. TRADITIONAL PERFECT FORESIGHT MODEL SOLUTION METHODS

Let $x_t \in R^n$ be a list of time t values for economic variables such as consumption, labor supply, capital stock, output, prices, interest rates, wages, etc., and z_t a list of exogenous variables, such as productivity levels, tax rates, monetary growth rates, etc., at time t. Perfect foresight models have the form

$$\begin{aligned} g(t, \overrightarrow{x}, \overrightarrow{z}) &= 0, \ t = 0, 1, 2, ... && (1) \\ x_{0,i} &= \overline{x}_{0,i}, \ i = 1, 2, ..., n_I && (2) \\ x_t & \quad \text{bounded} && (3) \end{aligned}$$

where

$$\begin{aligned} \overrightarrow{x} &\equiv (x_0, x_1, x_2, ..., x_s, ...) \\ \overrightarrow{z} &\equiv (z_0, z_1, z_2, ..., z_s, ...) \end{aligned}$$

and $g(t, \overrightarrow{x}, \overrightarrow{z}) \in R^n$ is a system of n functions concerning supply, demand, expectations, or other equilibrium relations among the economic variables. The equations in (1) represent Euler equations, market clearing conditions, and any other equations in the definition of equilibrium. Some of the economic variables may have fixed predetermined values at $t = 1$. These initial conditions are represented by the $n_I < n$ conditions in (2). The objective is to find a sequence of values for x_t such that (1,2) hold.

We shall use a simple example to illustrate our analysis. We examine the optimal growth problem

$$\begin{aligned} \max_{c_t} \quad & \sum_{t=0}^{\infty} \beta^t u(c_t) && (4) \\ s.t. \quad & k_{t+1} = F(k_t) - c_t \\ & k_0 = \overline{k}_0 \end{aligned}$$

The solution to this problem is the solution to the Euler equations

$$u'(c_t) = \beta u'(c_{t+1}) F'(k_{t+1}).$$

In the notation of (1,2), the growth problem can be expressed as

$$\begin{aligned} g_1(t, \overrightarrow{c}, \overrightarrow{k}) &\equiv u'(c_t) - \beta u'(c_{t+1}) F'(k_{t+1}) = 0,\ t = 0, 1, 2, ... \\ g_2(t, \overrightarrow{c}, \overrightarrow{k}) &\equiv k_{t+1} - F(k_t) + c_t = 0,\ t = 1, 2, ... \\ k_0 &= \overline{k}_0 \end{aligned} \tag{5}$$

In this problem, $x \equiv (c, k)$ and there are no exogenous variables. The functions in $g(t, \overrightarrow{c}, \overrightarrow{k})$ are the time t Euler equation and time t savings equation. The capital stock has a predetermined value at $t = 0$ but consumption is free at all times.The equations in (5) form a first-order nonlinear system in two variables. We can eliminate c_t and formulate the solution in terms of k_t only in the second-order nonlinear system

$$\begin{aligned} g(t, \overrightarrow{c}, \overrightarrow{k}) &\equiv u'(F(k_t) - k_{t+1}) - \beta u'(F(k_{t+1}) - k_{t+2}) F'(k_{t+1}) = 0,\ t = 0, 1, ... \\ k_0 &= \overline{k}_0 \\ \lim_{t \to \infty} k_t &\to k^{ss} \end{aligned} \tag{6}$$

We shall use (6) below as an example.

The system (1,2,3) is an infinite set of equations with an infinite number of unknowns. Under some conditions, there will be a locally unique solution; we will make that assumption as is implicitly done by all other methods. The problem in (4) has a unique solution for any given initial capital stock k_0. Any solution method must reduce the problem in some way. There are several ways to do this.

Most methods use *domain truncation* to reduce the problem to a finite-horizon problem. That is, they solve the truncated problem

$$g(t, x_0, x_1, ..., x_T, x^{ss}, x^{ss}, ..., \overrightarrow{z}) = 0,\ t = 0, 1, 2, ..., T-1 \tag{7}$$

$$x_{0,i} = \overline{x}_{0,i},\ i = 1, 2, ..., n_I \tag{8}$$

where x^{ss} is the steady state value, or some proxy for the long run. Some components of x_T are also fixed at their long-run values to make the number of unknowns in (7,8) equal to the number of equations. Domain truncation reduces (1,2,3) to (7,8), a system of nT nonlinear equations in nT unknowns. There is no boundedness equation in (7,8) since (7) imposes $x_t = x^{ss}$ for $t > T$. The long-run proxy x^{ss} need not be the steady state. It is taken to be some convenient value. The objective is that the choice of x^{ss} should not affect the solution since it lies in the distant future.

In (5), it would be natural to set c_t and k_t equal to their long-run steady state values for $t > T$. This would leave us with one too many unknowns. We could set c_T equal to its steady state value and leave k_T free, creating a system where there were equal number of c and k unknowns.

There is always the question of what T should be. Any method should try alternative values for T and accept a solution only when the choice of T does not substantially affect the solution. There are difficulties with this approach. For example, Kehoe has shown that the solution can be very sensitive to the choice of T, settling down only for very large values of T.

Since T is typically large, we need to develop special methods. Fortunately, we can apply methods from the literature on solving large systems (see, e.g., Kelley (1995), Saad (1996), Young (1971)). Some algorithms break the problem into smaller systems and then use solve individual systems iteratively until the full system converge. For example, Fair and Taylor(1983) use a block Gauss-Seidel procedure period by period for a given time choice of T and then tests the sensitivity to T. Other examples of this approach are Hall (1985), Fisher et al. (1986), Fisher (1992), and Hughes Hallett and Piscitelli (1998). Convergence of such methods depends on the order of the equations and is linear at best. The advantages are their simplicity and small memory requirements. However, they may not converge even after using various strategies including reordering of equations and damping factors.

An alternative is to use Newton's method. This is possible when the Jacobian is sparse. Juillard et al.(1998) pursued this strategy. Using Newton's method is difficult because the Jacobian is large. Gilli and Pauletto (1998) economize on this by using a Newton-style method together with a Krylov method to compute approximate Newton steps instead of exactly solving for the Newton step.

2. PARAMETRIC PATH METHOD FOR PERFECT FORESIGHT MODELS

The parametric path approach employs a substantially different strategy. Instead of treating each value of x_t as independent it applies some a priori knowledge about how x_t evolves over time. For example, the sequence 1,2,1,2,.. is not likely to represent a quarterly series for the capital stock or even aggregate consumption. Since the capital stock cannot change quickly, relative changes will be small over a small period of time. Similarly, consumption smoothing on the part of consumers imply that consumption paths will be smooth in the absence of unanticipated shocks. This feature of the solutions is not exploited by standard methods since they treat each distinct x_t separately. Instead, our intuition says that $\overrightarrow{x}$ should be a smooth function of time t. This insight allows us to dramatically reduce (1,2) to a much smaller system to which we can apply methods which could never be used directly on (1,2).

2.1. A Simple Parameterization. The key idea behind the parametric path method is to replace the sequence $\overrightarrow{x}$ with a parameterization of components of $\overrightarrow{x}$ in some way representing our beliefs that x evolves smoothly. First consider the

functional form

$$x_{t,i} = \Phi(t; a^i, \lambda) \equiv \left(\sum_{j=0}^{m} a_j^i t^j \right) e^{-\lambda t} + x_i^{ss} \left(1 - e^{-\lambda t} \right) \tag{9}$$

In (9) the matrix of coefficients

$$a = (a_j^i)_{j=0,..,m}^{i=1,..,n} \in R^{m \times n}$$

parameterize several polynomials in t. The extreme values of the path $x_i(t; a, \lambda)$ are

$$\begin{aligned} \Phi(0; a^i, \lambda) &= a_0^i \\ \Phi(\infty; a^i, \lambda) &= x^{ss} \end{aligned}$$

For any coefficient matrix a, the path in (9) is a convex combination of the initial value a_0 and the long-run value x^{ss} where the time-dependent weights are $e^{-\lambda t}$ and $1 - e^{-\lambda t}$. If $a_j = 0$ for $j = 1, 2, ...$, then the path in (9) converges smoothly and monotonically to the x^{ss}. Since the exponential term dominates any polynomial term in (9), the path in (9) asymptotically satisfies the linear adjustment process

$$\frac{d}{dt} (x - x^{ss}) = -\lambda (x - x^{ss}) \tag{10}$$

The initial conditions (2) imply that

$$a_0^i = \overline{x}_{0,i}, \ i = 1, 2, ..., n_I \tag{11}$$

Other than the conditions in (11), we are free to choose the components of a so that the sequence

$$\overrightarrow{x} = \left(\begin{pmatrix} \Phi_1(0; a, \lambda) \\ \vdots \end{pmatrix}, \begin{pmatrix} \Phi_1(1; a, \lambda) \\ \vdots \end{pmatrix}, \begin{pmatrix} \Phi_1(2; a, \lambda) \\ \vdots \end{pmatrix}, ... \right)$$

approximately solves (1).

Some features of (9) can be determined by a priori knowledge about (1) The form in (9) imposes convergence to x^{ss} in the long run. Since we often know the steady state x^{ss}, we can use this knowledge in (9). The choice of λ is also one which can use a priori information. Given the asymptotic behavior in (10), we want λ to be the rate of convergence associated with the dominant eigenvalue of the linearization of (1) around x^{ss}. Sometimes we can compute this, but we often have good guesses about the asymptotic rate of convergence. We will see that a good guess is satisfactory. Otherwise, (9) is quite flexible for the initial values of t. Therefore, (9) is a form

which reflects our knowledge about the long-run behavior of the long-run equilibrium and our comparative ignorance about equilibrium in the short run.

The formula in (9) treats t as a continuous variable. This may initially seem odd to do in a discrete-time framework. However, there is nothing in (1) which requires t to be an integer. For example, the expression for the time t equation

$$u'\left(F(k_t)-k_{t+1}\right)-\beta u'\left(F(k_{t+1})-k_{t+2}\right)F'(k_{t+1}) \tag{12}$$

in (6) makes perfectly good sense for noninteger t if we define $k_t = x(t;a,\lambda)$ for some a and λ. We shall proceed as if t were continuous, making it an integer only when necessary. This is key since the whole idea of the parametric path method is to apply approximation methods for functions of a continuous variable to functions on the integers.

2.2. General Parameterizations. The general idea is to find some functional form[2] $\Phi(t;a)$ and use it to approximate the sequence as in

$$x_{t,i} = \Phi(t;a^i),\ i=1,..,n \tag{13}$$

The parametric path method is an application of projection methods (see Judd (1992)) for solving functional equations. One natural choice for our parameterization[3] is to use orthogonal polynomials of the form

$$\Phi(t;a) = \left(\sum_{j=0}^{m} a_j\phi_j(t)\right)e^{-\lambda t} + x^{ss}\left(1-e^{-\lambda t}\right) \tag{14}$$

where the family $\left\{\phi_j(t)\right\}_{j=0}^{\infty}$ is orthogonal with respect to $e^{-2\lambda t}$. More specifically, we assume

$$\phi_j(t) = L_j(2\lambda t)e^{-\lambda t} \tag{15}$$

where $L_j(s)$ is the degree j Laguerre polynomial. Laguerre polynomials are defined by the recursive formulas

$$L_{m+1}(x) = \left(\frac{2m+1-x}{m+1}\right)\ L_m(x) - \frac{m}{m+1}\ L_{m-1}(x)$$

The key property of the ϕ_i is that they are mutually orthogonal; that is

$$\int_0^{\infty}\phi_j(t)\phi_i(t)dt = \int L_j(2\lambda t)e^{-\lambda t}L_i(2\lambda t)e^{-\lambda t} = 0,\ i\neq j$$

[2]Here we use the notation $\Phi(t;a)$ and do not include λ explicitly. If there were a parameter in (13) like the λ parameter in (9), we fold it into the parameter list a.

[3]See Judd(1992) and Judd(1998) for discussions of orthogonal polynomials and their use in developing projection method algorithms for dynamic general equilibrium methods.

because the Laguerre polynomials $L_i(2\lambda t)$ are orthogonal with respect to the weight $e^{-2\lambda t}$. Since the functional form in (14) is a linear combination of ϕ_i functions, we have an orthogonal representation of functions of the form (9). The space of functions spanned by the form (14) is the same as that spanned by the form (9). The form in (14) is also flexible for initial values of t and imposes convergence to x^{ss} in the long run. We prefer the orthogonal representation in (14) since, as we will see, it has many computational advantages.

Of course, there are many possible forms for the functional form $\Phi(t;a)$. Φ should be flexible enough to parsimoniously approximate any likely solution. The best choice of n cannot be determined *a priori*. Generally, the only correct choice is $n = \infty$. If the choice of the functional form is good then larger n will yield better approximations. We are most interested, however, in the smallest n that yields an acceptable approximation. We initially begin with small n and increase n until some diagnostic indicates little is gained by continuing. Computational considerations also play a role in choosing a functional form. Φ should be simple to compute and each coefficient should be of roughly the same importance.

Our task is somewhat simplified since each component of Φ is a one dimensional function of t. This fact implies that we can use the full range of possible functional forms, including splines and rational polynomials as well as polynomial systems. We will stay with polynomial systems for the purposes of this study.

2.3. Projection Conditions. Once we have chosen a parameterization Φ, we need to devise some way to choose the coefficients a in our approximation so that Φ approximates an equilibrium to (1,2). Define the residual function

$$R(t,a) = g(t, x(0;a), x(1;a), ..., x(s;a), ...)$$

where again we are treating t provisionally as a continuous variable. We want to find some a such that $R(t;a)$ is "practically" zero for all t,implying that equations (1,2) nearly hold. For such an a, the time path $x(t;a)$ will be an approximate solution to the perfect foresight model (1,2).

To proceed, we need to define what it means for $R(t;a)$ to be small. There are several ways to do this. The first direct way is to create for each choice of a the "sum of squared residuals" index

$$SSR(a) = \int_0^\infty R(t;a)^2 w(t) dt \tag{16}$$

where $w(t) > 0$ is some weighting function. Normally one would not want to choose $w(t)$ since the domain of integration is infinite. However, the functional form (9) converges to the true steady state. For that parameterization, $R(t;a)$ will go to zero exponentially for any choice of a, implying that (16) is well-defined even if $w(t) = 1$.

The general *least squares projection method* computes the L^2 norm of the residual function, $SSR(a)$, for some w and chooses a to solve

$$\min_a \ SSR(a). \tag{17}$$

This reduces the problem of solving an infinite number of equations to solving a nonlinear minimization problem in R^n, a tractable problem. Of course, the standard difficulties will arise. For example, there may be local minima which are not global solutions. However, there is no reason for these problems to arise more often here than in any other context, such as maximum likelihood estimation, where minimization problems are solved numerically.

The least squares method is a direct implementation of the idea to make small the error of the approximation. In general, one could develop alternative implementations by using different norms. However, most projection techniques find a good-fitting approximation in a less direct but more effective fashion. The key concept is that of a *projection*. Specifically, we choose a weight function $w(t)$, a set of test functions, $p_i(t)$, and form projections of the form

$$\langle R_i(t;a), p_j(t) \rangle \equiv \int_0^\infty R_i(t;a)\ p_j(t)\ w(t)\ dt \tag{18}$$

For these techniques the basic idea is that the true solution would produce a zero residual error function; in particular, the residual function would have a zero projection in all directions. Therefore one way to find the nm components of a is to fix n projections and choose a so that the projection of the resulting residual function in each of those n directions is zero. That is, we want to find a such that

$$\langle R_i(t;a), p(t) \rangle = 0,\ i = 1, .., n \tag{19}$$

for several test functions $p\,(t)$. The equations in (19) are not sufficient since the initial conditions must also be included in the analysis. The initial conditions (2) imply

$$x_{0,i}(0) = \sum_{j=0}^{m} a_j^i \phi_j(0) = \overline{x}_{1,i},\ i = 1, 2, ..., n_I \tag{20}$$

We want to find a which satisfy (20) as well as several conditions of the form (19).

2.4. Integral Approximations. The integral in (18) needs to be computed approximately. At this point, we need to drop the fiction of t being continuous. We will use standard integration methods to motivate the integral approximations. Integration methods generally apply formulas of the form

$$\int_0^\infty h(t)w(t)dt \doteq \sum_{\ell=1}^{m} \omega_\ell h(t_\ell)$$

for some quadrature weights ω_i and quadrature nodes t_i. The quadrature nodes t_i will typically not be integers. Since the residual function $R(t;a)$ for (1,2) is only really defined at integer t we use the approximation

$$\int_0^\infty h(t)w(t)dt \doteq \sum_{\ell=1}^{m} \omega_\ell h(t_\ell^I) \tag{21}$$

where t_i^I is the nearest integer to t_i.

It is only at this point, the final detail in the algorithm, that we need to restrict t to be an integer. It is the integration formula in (21) which tells us which t's to use. The maximum t tells us how much into the future we need to approximate the x_t path. Fair-Taylor and other traditional methods make an ad hoc determination of the horizon they use. Also, the t's which are used depend on the integration formula we use in (21). In fact, the t's we use are often optimal given the integration formula used.

In the end, the parametric path method uses the integral approximation (21) and reduces (1,2) to solving the system of nonlinear equations

$$P_{ij}(a) \equiv \sum_{\ell=1}^{m} \omega_\ell R(t_\ell^I;a)p_j(t_\ell^I) = 0,\ i = 1,..,n,\ j = 0,..,m \tag{22}$$

$$\Phi(0;a) = \overline{x}_{1,i},\ i = 1,2,...,n_I \tag{23}$$

The system (22,23) is overidentified. We want to impose (23), so we drop n_I equations from (22); it is normally preferable to drop the projections from some higher-order polynomial test functions. When we refer to the system (22,23) we will implicitly be referring to an exactly identified system.

Here we see a critical feature of the parametric path method. In the end, we evaluate the equilibrium equations at only a small number of times t. Other methods compute (1) for all $t < T$ for some large T.

Different choices of the p_i defines different implementations of the projection method. We will use the *Galerkin* method, also known as the *Bubnov-Galerkin* or *Galerkin-Petrov* method. In the Galerkin method the test functions are the basis functions in a linear representation and the weighting function is chosen so that the basis functions are mutually orthogonal. This produces a projections of the form

$$P_i(a) \equiv \langle R(x;a),\ \varphi_i(x) \rangle = 0,\ i = 1, \cdots, n.$$

The Galerkin method is just one possible alternative. Most of the others described in Judd (1992) could also be used.

2.5. Initial Guesses. Good initial guesses are important since projection methods involve either a system of nonlinear equations or optimizing a nonlinear objective. One advantage of the parametric path method is that there is a very natural initial guess. We know the steady state values for all variables, and we often have a good guess for the asymptotic rate of convergence to the steady state. A natural initial guess is the path which smoothly moves from the initial state x_0 to the steady state x^{ss} at the asymptotic rate of convergence λ :

$$x^{init}(t) = x_0 e^{-\lambda t} + x^{ss}\left(1 - e^{-\lambda t}\right) \tag{24}$$

There is a general approach which is often useful. The least squares approach may not be a good one to use for high-quality approximations. However, it may yield low-quality approximations relatively quickly, and, since the least squares method is an optimization method, convergence to a local extrema is ensured even if one has no good initial guess. Furthermore, by adding terms to the least squares objective, one can impose sensible restrictions on the coefficients to eliminate economically nonsensical extrema. These facts motivate a two-stage approach. First, one uses a least squares approach with a loose convergence criterion to quickly compute a low-quality approximation. Second, one uses this approximation as the initial guess for a projection method attempting to compute a higher-order approximation. With some luck the least squares solution will be a good initial guess for the second computation.

2.6. Finding the Solution. To identify the coefficients a we either use a minimization algorithm to solve (17) or a nonlinear algebraic equation solver to solve (22,23). We could create an overidentified system $P(a) = 0$ and then use a nonlinear least squares algorithm to find the least squares solution. The nonlinear equations associated with Galerkin and other inner product methods can be solved by the variety of nonlinear equation methods. While fixed-point iteration appears to be popular in economics, Newton's methods and its refinements have often been successful. A main advantage of parametric path methods is that it reduces the problem to a small dense nonlinear equation system to which Newton and similar methods can be effectively applied.

If other methods do not work and it is difficult to find good initial guesses, then one can use globally convergent homotopy methods since they do not require good initial guesses. Homotopy methods are not possible to apply to conventional formulations since homotopy methods are intractable when the number of unknowns is large.

2.7. Checking the Solution. The system (22,23) uses (1) only at only a small number of t's. Before we accept it, we need to check any solution to (22,23) at some of the t's we did not use. Suppose that the solution to (22,23) implies the

approximation $x_t = \Phi(t; a)$. To make the parametric path method comparable to traditional approaches, we evaluate

$$E = \max_{t=0,1,\ldots,T} \|g(t, \overrightarrow{x}, \overrightarrow{z})\|$$

where $\overrightarrow{x}$ is the time path implied by $x_t = \Phi(t; a)$ and T is some large time. The index E is the maximum error in the equations (1) over a long range of time. Any traditional method continues until E is small; this test is really part of any conventional stopping criterion for solving nonlinear systems. This check allows us to use the same stopping rule as traditional methods. In the parametric path method, if a solution to (22) does not imply a sufficiently small value for E we can begin again with a more flexible parameterization. In the case of an orthogonal polynomial approximation this means using higher-order polynomial terms.

We summarize the parametric path method.

Parametric Path Method: Summary	
Step 1:	Choose parameterization $x = \Phi(t; a)$.
Step 2:	Form residual function $R(t; a) = g(t, \Phi(.; a), z)$.
Step 3:	Select test functions $p_j(t)$.
Step 4:	Form projections $P_{ij}(a) \doteq \langle R_i(t; a), p_j(t) \rangle$,using integration formulas where necessary.
Step 5:	Solve system of $P_{ij}(a) = 0$ equations plus initial conditions.
Step 6:	Compute $E = \max_{t=0,1,\ldots,T} \|g(t, \overrightarrow{x}, \overrightarrow{z})\|$; accept a if E is sufficiently small; otherwise begin again at Step 1 with a more flexible approximation.

3. GROWTH EXAMPLE

We now apply the parametric path method to our optimal growth problem displayed in (4). We will solve the system (6) While this example is a very simple one, it does share the essential features present in many other perfect foresight models: the system is sparse and nearly diagonal, and equilibrium converges linearly to the steady state.

3.1. Parameterization. We know from theory that the path of capital which solves (6) converges asymptotically to the steady state at a linear rate equal to λ where λ is the stable eigenvalue of the linearized system around the steady state k^{ss}. We have the initial condition $k(0) = k_0$. We also know that the time path of capital is "smooth" in that convergence is monotone. This suggests the parameterization

$$k(t) = \left(k_0 + \sum_{j=1}^{m} a_j t^j e^{-\lambda t}\right) + k^{ss}\left(1 - e^{-\lambda t}\right) \tag{25}$$

We want to express the parameterization in an orthogonal fashion since orthogonal parameterizations have better numerical properties. We need to construct an orthogonal family which spans the same space as that in (25). Let

$$\phi_j(t) = L_j(2\lambda t)e^{-\lambda t}$$

where $L_j(t)$ is the degree j Laguerre polynomial. Laguerre polynomials are defined by the recursive formulas

$$L_{n+1}(t) = \frac{1}{n+1}\ (2n+1-t)\ L_n(t) - \frac{n}{n+1}\ L_{n-1}(t) \tag{26}$$

and are satisfy $\int_0^\infty L_i(t)L_j(t)e^{-t}dt = 0$ for $i \neq j$. Therefore,

$$\int_0^\infty \phi_i(t)\phi_j(t)dt = \int_0^\infty L_i(2\lambda t)L_j(2\lambda t)e^{-2\lambda t}dt = 0,\ i \neq j$$

We use the orthogonal parameterization

$$k(t) = \sum_{j=0}^{m} a_j\phi_j(t) + k^{ss}\left(1 - e^{-\lambda t}\right) \tag{27}$$

The initial condition is

$$k(0) = k_0 = \sum_{j=0}^{m} a_j\phi_j(0)$$

and implies

$$a_0 = \phi_0(0)^{-1}\left(k_0 - \sum_{j=1}^{m} a_j\phi_j(0)\right)$$

Therefore, the parameterization is

$$k(t) = \sum_{j=1}^{m} a_j\phi_j(t) + \phi_0(0)^{-1}\left(k_0 - \sum_{j=1}^{m} a_j\phi_j(0)\right) + k^{ss}\left(1 - e^{-\lambda t}\right) \tag{28}$$

and the unknowns are the coefficients $a_i,\ i = 1, 2, ..., m$.

There are two key features of (28). First, the exponential decay terms $e^{-\lambda t}$,some of which are in the ϕ terms and some explicitly in (28), impose the boundedness conditions. Second, the initial conditions are also satisfied for any choice of $a \in R^n$. These facts allow us to focus on finding an a which produces a good solution to (6).

3.2. Projection Conditions. We define the residual function

$$R(t,a) = u'\left(F(k(t)) - k(t+1)\right) - \beta u'\left(F(k(t+1)) - k(t+2)\right) F'(k(t+1))$$

Note that $R(t;a)$ is well-defined for any real value of t, not just the integers, since $k(t)$ is defined for all t in (28). This observation makes us more comfortable with provisionally using continuous variable methods. We want to choose a so that $R(t;a)$ is nearly zero for all t. By construction, $R(t;a) \to 0$ and $t \to \infty$. The coefficients a adjust $k(t)$ so that $R(t;a)$ is small for finite t. We will use a Galerkin method where the test functions are the basis functions. To this end, we define the set of projection formulas

$$P_i(a) = \int_0^\infty R(t;a) L_j(2\lambda t) e^{-\lambda t} dt, \ i = 0, 1, ... \tag{29}$$

The key problem is computing the integrals in (29).

We will use a change of variable (COV) method[4]. Specifically, we will use the exponential and logistic maps

$$\begin{aligned} t &= -\frac{1}{L}\log\left(\frac{1-x}{2}\right) \\ x &= 1 - 2e^{-Lt} \end{aligned}$$

which will map $x \in [-1,1]$ to $t \in [0,\infty)$, and vice versa. The projection equations are then transformed into

$$\begin{aligned} P_i(a) &= \int_0^\infty R(t;a)\phi_i(t)dt \\ &= \int_{-1}^1 R(t(x);a)\phi_i(t(x))t'(x)dx \end{aligned}$$

We next use Gauss-Chebyshev integration formula to arrive at

$$P_i(a) = \frac{\pi}{n}\sum_{\ell=1}^n R(t(x_\ell);a)\phi(t(x_\ell))(1-x_\ell^2)^{1/2}t'(x_\ell)$$

where the Chebyshev integration nodes are

$$x_\ell = \cos\left(\frac{2\ell+1}{2n}\pi\right)$$

[4]We could use any of a large number of integration methods. For example, Gauss-Laguerre integration would be natural because of the presence of the $e^{-\lambda t}$ term in (29). We chose to use a COV method because there is a bit more flexibility and because transforming the problem to $[-1,1]$ allows us to use equioscillation ideas to evaluate the error.

However, we really should use just the integers, so the final form for each projection[5] is

$$P_i(a) = \sum_{\ell=1}^{n} R(t^I(x_\ell); a)\phi_i(t^I(x_\ell))(1 - x_\ell^2)^{1/2} t'(x_\ell)$$

where $t^I(x)$ is the value of $t(x)$ rounded to the nearest integer.

Since $a = (a_1, a_2, ..., a_m) \in R^m$, we solve the system of equations

$$P_j(a) = 0, \ j = 0, 1, ..., m-1 \tag{30}$$

3.3. Numerical Results. We apply the parametric path method to (4) assuming $u(c) = c^{1+\gamma}/(1+\gamma)$ and $F(k) = k + Ak^\alpha$ where we choose A always so that the steady state is $k^{ss} = 1$. We display the results for difficult cases. If the initial capital stock k_0 is close to the steady state then the exponential component is a very good approximation without any help from polynomial terms. Therefore, we choose $k_0 = .5$. We examine $\gamma = -0.5, -2.0, -5$ 0. We assume $\alpha = .25$; α parameter had little effect on algorithm performance. We choose $\beta = .99$, essentially modelling a three-month period of time. We used a Powell hybrid method for solving (30) and the initial guess set all polynomial terms equal to zero.

We first computed the optimal consumption policy function $C(k)$ for (4) using the methods outlined in Judd (1992). The Euler equation errors from that method were on the order of 10^{-6} so we took that to be the truth. We used $C(k)$ to compute the true path for k and c. These paths were then used to assess the accuracy of the parametric path method.

Tables 1, 2, and 3 report the maximum relative error in the k path where the maximum is taken over the first 2500 periods. That is, we compute the true path k_t, the path $\widehat{k}_t$ produced by applying the parametric path method, and report the maximum error

$$E = \max_{t=1,..,2500} \frac{\left|k_t - \widehat{k}_t\right|}{k_t}$$

A key step in this example of the parametric path method is choosing λ in (28). The asymptotic rate of convergence to the steady state of (4) is a natural choice; let μ be that value. It is easy to compute μ for each parametric case of (4) we examine, so we use μ in Table 1. Table 1 examines several choices of m, the degree of the polynomial pieces of our approximation (15). Table 1 shows that the method does quite well. Fourth-degree approximations uniformly have errors of at most .02%, implying essentially four-digit accuracy. We report the errors in k since the errors in the consumption path are less than half the errors in k. The errors drop as we increase the flexibility in the parameterization, dropping by about half for each extra parameter. Also, the relative errors are nearly the same across different values of γ.

[5] We also drop the π/n factor since we want to set $P = 0$.

Table 1: Maximum errors in $k : \lambda = \mu$

	m :				
γ	1	2	3	4	5
-5.0	1(-3)	6(-4)	3(-4)	2(-4)	1(-4)
-1.1	2(-3)	7(-4)	3(-4)	2(-4)	1(-4)
-0.5	4(-3)	1(-3)	6(-4)	3(-4)	2(-4)

It is not surprising that the results in Table 1 are so good since the correct choice for λ nails the approximation asymptotically. We next show that the parametric path algorithm would also work if we don't have a good guess for λ. Table reports the same exercise except we take $\lambda = 3\mu$ in each case, and Table 3 chooses $\lambda = .33\mu$. The maximum errors are greater, but also drops as we add parameters.

Table 2: Maximum errors in $k : \lambda = 3\mu$

	m :				
γ	1	2	3	4	5
-5.0	1(-1)	5(-2)	2(-2)	1(-2)	6(-3)
-1.1	1(-1)	6(-2)	3(-2)	1(-2)	7(-3)
-0.5	1(-1)	6(-2)	3(-2)	1(-2)	7(-3)

Table 3: Maximum errors in $k : \lambda = .33\mu$

	m :				
γ	1	2	3	4	5
-5.0	1(-3)	4(-2)	1(-2)	6(-3)	2(-3)
-1.1	1(-1)	4(-2)	2(-2)	9(-3)	1(-3)
-0.5	1(-1)	4(-2)	2(-2)	7(-3)	2(-3)

Some other indices indicate the efficiency of the parametric path algorithm. We used Newton's method and it converged in 3 or fewer iterations, even in the cases where λ was different from μ. Also, the orthogonal nature of the approximation makes the Jacobian well-behaved. The condition number of the Jacobians were all small. Furthermore, the Jacobians were nearly triangular with the diagonal elements dominating the elements below the diagonal.

4. Comparisons with Alternative Methods

The parametric path method initially appears to be very different from standard methods. The differences are less than they seem. We next compare it with alternative methods in order to highlight the common features. These comparisons will also indicate how they can be combined in hybrid methods which take advantage of each methods' strengths.

The problem (1) is an infinite system of equations in an infinite number of unknowns. The solution lies in R^∞. It is impossible to solve such an infinite system on a computer. We need to reduce the dimensionality somehow. The conventional approach finds an approximation which lies in the finite-dimensional subspace

$$X_T^{FT} = \{x \in R^\infty | x_t = x^{ss},\ t > T\}.$$

Successively better approximations are produced by increasing the terminal date T. We also reduce the problem (1) to a finite-dimensional space. The space we used in our example was

$$X_{m,\lambda}^{PP} = \left\{ x \in R^\infty | x_t = e^{-\lambda t} \sum_{i=0}^{m} a_i t^i \right\}.$$

In the limit as we take T or m to infinity, both finite-dimensional spaces span R^∞. The issue is which approach produces small finite-dimensional spaces which approximates well the true solution.

The main idea behind the parametric path method is similar in spirit to Krylov methods (and their early forms, such as conjugate gradient methods). The inner loop of a Krylov method reduces a large problem to a smaller one which generates an approximation using ideas similar to our projection method. The outer loop examines a succession of finite-dimensional approximations, where the new directions are chosen to keep the smaller finite-dimensional problems well-behaved. The parametric path method also continues by examining successively larger approximation spaces until the apparent error is small. In our version of the parametric path method, the sequence of spaces used is exogenously specified, but a more refined version could endogenize the sequence of approximations used.

The projection equations system (22) could also be a large system. We noted that (22) was nearly triangular in our example. This property is actually expected given the orthogonality of the projections and indicates that Gauss-Seidel methods, such as those used in Hughes Hallet and Piscitelli could be applied. If we used a finite-element approach, (22) would be sparse and could use the Newton style methods used in Juillard et al. and Gilli and Pauletto. The main accomplishment of the parametric path method is the reduction in dimensionality. The reduced system can still take advantage of many other techniques for solving large systems.

Some systems are not sparse. In particular, the overlapping generations analysis of Auerbach and Kotlikoff produces systems of equations which are too dense for sparse methods. They use a successive approximation procedure to compute equilibrium. However, at best it displays on linear convergence and sometimes it does not converge. The parametric path method does not exploit sparseness and could be also used to solve models such as Auerbach and Kotlikoff. The key assumption is that the economic variables behave in a relatively smooth fashion and that a low-dimensional approximation is good once one focuses the search on a suitable space of solutions.

The parametric path method can be useful even in cases where the equilibrium paths of economic variables are not smooth. The projection aspect of the parametric path method implies that the parametric path method will likely produce a smoothed approximation of the true path. This can then be used as an initial guess for a more refined method such as Fair-Taylor or any of the other later methods.

5. CONCLUSION

The parametric path method develops a new approach to solving perfect-foresight models. It parsimoniously parameterizes the time path of the unknown economic variables. The parameterization allows us to apply ideas from approximation theory to the problem of determining the undetermined coefficients. In particular, the use of orthogonal polynomials allows us to use Gaussian integration methods which tell us which equations to use in approximating the solution. The algorithm is more flexible than most alternatives. For example, it can be used even for non-sparse systems, such as overlapping generations models. It is also more robust; the reduction in the number of unknowns allows us to use more reliable nonlinear equation solution methods.

REFERENCES

[1] Boucekkine, R., 1995. An Alternative Methodology for Solving Nonlinear Forward-Looking Models. *Journal of Economic Dynamics and Control* 19 (4), 711-734.

[2] Fair, R.C., and Taylor, J.B., 1983. Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models. *Econometrica* 51 (4), 1169-1185.

[3] Fisher, P.G., and Hughes Hallett, A.J., 1987. The convergence characteristics of iterative techniques for solving econometric models. *Oxford Bulletin of Economics and Statistics* 49, 231- 244.

[4] Fisher, P.G., and Hughes Hallett, A.J., 1988a. Iterative techniques for solving simultaneous equation systems. a view from the economics literature. *Journal of Computational and Applied Mathematics* 24, 241- 255.

[5] Gagnon, J.E., 1991. A Forward-Looking Multicountry Model for Policy Analysis: MX3. *Economic and Financial Computing* 1, 311- 361.

[6] Gilli, M, and Pauletto, G, 1997. Sparse direct methods for model simulation. *Journal of Economic Dynamics & Control* 21, 1093- 1111.

[7] Gilli, M., and Pauletto, G., 1998. Nonstationary iterative methods for solving models with forward looking variables. *Journal of Economic Dynamics and Control*, this issue.

[8] Hughes Hallett, A., and Piscitelli, L., 1998. A new convergence theorem for successive overrelaxation iterations, *Computational Economics*.

[9] Hughes Hallett, A.J., and Piscitelli, L., 1998. Simple reordering techniques for expanding the convergence radius of first order iterative techniques. *Journal of Economic Dynamics & Control*.

[10] Juillard, M., Laxton, D., McAdam, P., and Pioro, H., 1998. An algorithm competition: first-order iterations versus Newton-based techniques. *Journal of Economic Dynamics & Control*.

[11] Kelley, C.T, 1995. *Iterative Methods for Linear and Nonlinear Equations*. SIAM, Philadelphia.

[12] Pauletto, G., 1995. Solution and simulation of macroeconometric models, Thesis, Department of Econometrics, University of Geneva, Switzerland.

[13] Pauletto, G, 1997. *Computational Solution of Large-Scale Macroeconometric Models. Series: Advances in Computational Economics*. Kluwer Academic Publishers, Dordrecht.

[14] Saad, Y., 1996. Iterative Methods for Sparse Linear Systems. PWS Publishing Comp., Boston, MA.9

[15] Saad, Y, and Schultz, M, 1986. GMRES: a generalized minimal residual algorithm for solving nonsymetric linear systems. *SIAM Journal for Scientific and Statistical Computing* 7, 856- 869.

[16] Young, D.M., 1971. *Iterative Solution of Large Linear Systems*. Academic Press, New York.

6. EXTRA STUFF

Rational Chebyshev functions on the semi-infinite interval

$$TL_n(t) = T_n\left(\frac{t-L}{t-L???}\right)$$

pseudospectral grid points

$$y_i = L\cot(x_i/2), x_i = \pi\frac{2i-1}{2N}, i = 1, ..., N$$

Boyd(1982a, J Comp Phys)

Boyd(1982a, b, "Stieltjes" Math Comp) for guidance on L

Exponential map asymptotically inferior, but sometimes enjoys finite-expansion advantages. It is, however, dangerous. P. R. Spalart(1984, Contemp. Math.)

[see Boyd, p. 408ff, semi-inf chapter]

$$\begin{aligned} R(t;a) &\sim e^{-\lambda t} \\ f'(k) - f'(k^{ss}), R_a &\sim e^{-\lambda t} \end{aligned}$$

$$\begin{aligned} R(t,a) &= f(k(t;a)) = 0 \\ R_{a_j} &= f'(k)\left(\phi_j - e^{-\lambda t}\right) = f'(k)(L_j(2\lambda t) - 1)e^{-\lambda t} \end{aligned}$$

We need to define the concept of an approximate solution.

$$\begin{aligned} P^j(a) &= \int_0^\infty R(t;a)L_j(2\lambda t)e^{-\lambda t}dt \\ P_i^j(a) &= \int_0^\infty R_{a_i}(t;a)L_j(2\lambda t)e^{-\lambda t}dt \\ &= \int_0^\infty f'(k)(L_i(2\lambda t) - 1)e^{-\lambda t}L_j(2\lambda t)e^{-\lambda t}dt \end{aligned}$$

problem: we impose a false value on $k(T)$.

It is convenient to define some function $t(\zeta) : [0,1] \to [0,\infty)$ which will allow us to use methods for approximating functions on $[-1,1]$ for approximating functions on $[0,\infty)$. Specifically, we could use

$$\begin{aligned} \zeta(t) &= 1 - 2e^{-Lt} \\ t(\zeta) &= -L\ \log\left(\frac{1-\zeta}{2}\right) \end{aligned}$$

For such a COV function, we could let

$$\Phi(t;a) = \Psi(\zeta(t);a)$$

where $\Psi(\zeta;a) : [-1,1] \rightarrow R$ is a functional form mapping $[-1,1]$ to the real line. Another possible form for Ψ is the spline form

$$\Psi(\zeta;a) =$$

Both splines and polynomials are linear in most of their unknown coefficients. One could also use nonlinear function forms such as rational Chebyshev functions

$$\begin{aligned} \zeta(t) &= ?? \\ t(\zeta) &= T_j\left(\frac{1+\zeta}{1-\zeta}\right) \end{aligned}$$

or neural networks

$$\Psi(\zeta;a) =$$

Alternative choices for a COV function are

$$\begin{aligned} \zeta(t) &= ?? \\ t(\zeta) &= L\,\frac{1+\zeta}{1-\zeta} \end{aligned}$$

When we have chosen a COV $t(\zeta)$, the problem ?? reduces to finding

Gauss-Laguerre

$$\begin{aligned} E(a) &= \int_0^\infty R(t;a)\phi(t)dt \\ &= \int_0^\infty R(t;a)\phi(t)e^{\lambda t}e^{-\lambda t}dt \\ &= \sum \omega_i R(t_i;a)\phi(t_i)e^{\lambda t} \\ &\doteq \sum \omega_i^I R(t_i^I;a)\phi(t_i^I)e^{\lambda t} \end{aligned}$$

COV methods

$$\begin{aligned} E(a) &= \int_0^\infty R(t;a)\phi(t)dt \equiv \int_0^\infty h(t)dt \\ &= \int_{-1}^1 h(t(x))t'(x)dx \\ &= \sum \omega_i h(t(x_i))t'(x_i) \\ &\doteq \sum \omega_i^I h(t^I(x_i))t'(x(t_i^I)) \end{aligned}$$

Exponential-COV-Gauss-Chebyshev

$$\begin{aligned} t &= -L\ \log(1-x) \\ t &\in [0,\infty),\ x \in [-1,1] \end{aligned}$$

Algebraic-COV-Gauss-Chebyshev

$$\begin{aligned} t &= L\,\frac{1+x}{1-x} \\ t &\in [0,\infty),\ x \in [-1,1] \end{aligned}$$

Rational Chebyshev functions on the semi-infinite interval??

$$TL_n(t) = T_n\left(\frac{t-L}{t-L}??\right)$$

pseudospectral grid points

$$y_i = L\cot(x_i/2), x_i = \pi\frac{2i-1}{2N}, i = 1, ..., N$$

Boyd(1982a, J Comp Phys)

Boyd(1982a, b, "Stieltjes" Math Comp) for guidance on L

Exponential map asymptotically inferior, but sometimes enjoys finite-expansion advantages. It is, however, dangerous. P. R. Spalart(1984, Contemp. Math.)

[see Boyd, p. 408ff, semi-inf chapter]As we have seen in our examples, projection techniques include a variety of special methods. In general, we specify some inner product, $\langle\cdot\,,\,\cdot\rangle$, of B_2, and use $\langle\cdot\,,\,\cdot\rangle$ to measure the "size" of the residual function, R, or its projection against the test functions. We can use inner products of the form

$$\langle f(x)\,,\,g(x)\rangle \equiv \int_D f(x)g(x)w(x)\,dx$$

for some weighting function $w(x)$, but there is no reason why we are limited to them. In choosing the norm, one should consider exactly what kind of error should be small and find a norm that will be sensitive to the important errors. There are several ways to proceed.

One approach would be to use *domain truncation* where we approximate the future with some distant future time:

$$\begin{aligned} P_j(a) &= \int_0^\infty R(t;a)\phi_j(t)dt \\ &\doteq \int_0^T R(t;a)\phi_j(t)e^{\lambda t}e^{-\lambda t}dt \\ &= \sum \omega_i R(t_i;a)\phi_j(t_i)e^{\lambda t} \\ &\doteq \sum \omega_i^I R(t_i^I;a)\phi_j(t_i^I)e^{\lambda t} \end{aligned}$$

and then approximate ?? in some fashion

$$\begin{aligned} P(a) &= \int_0^\infty R(t;a)\phi_j(t)dt \\ &\doteq \int_0^T R(t;a)\phi_j(t)e^{\lambda t}e^{-\lambda t}dt \\ &= \sum \omega_i R(t_i;a)\phi_j(t_i)e^{\lambda t} \\ &\doteq \sum \omega_i^I R(t_i^I;a)\phi_j(t_i^I)e^{\lambda t} \end{aligned}$$

problem: we impose a false value on $k(T)$.

One could use Gauss-Laguerre integration formulas???

$$\begin{aligned} P(a) &= \int_0^\infty R(t;a)\phi_j(t)??dt \\ &= \int_0^\infty R(t;a)\phi_j(t)e^{\lambda t}e^{-\lambda t}dt \\ &= \sum \omega_i R(t_i;a)\phi_j(t_i)e^{\lambda t} \\ &\doteq \sum \omega_i^I R(t_i^I;a)\phi_j(t_i^I)e^{\lambda t} \end{aligned}$$

$$\begin{aligned} P_j(a) &= 0,\ j = 0,1,...,m \\ P_i^j(a) &= \int_0^\infty R_{a_i}(t;a)L_j(2\lambda t)e^{-\lambda t}dt \\ &= \int_0^\infty f'(k)(L_i(2\lambda t)-1)e^{-\lambda t}L_j(2\lambda t)e^{-\lambda t}dt \end{aligned}$$

7. BIG REFERENCE LIST

REFERENCES

[1] Armstrong, J., Black, R., Laxton, D., and Rose, D., 1995. A robust method for simulating forward-looking models. *The Bank of Canada's New Quarterly Projection Model*, Part 2, Technical Report No. 73, Ottawa, Canada.

[2] Axelsson, O, 1985. Incomplete block matrix factorization preconditioning methods. The ultimate answer? *Journal of Computing Applied Mathematics* 12, 3- 18.

[3] Axelsson, O, 1994. *Iterative Solution Methods.* Oxford University Press, Oxford, UK.

[4] Bank of England, 1995. *Macroeconomic Forecasting Model*, Monetary Affairs. Bank of England, London.

[5] Barrett, R et al., 1994. Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods. SIAM, Philadelphia, PA.

[6] Bartolini, L., Razin, A., and Symansky, S., 1995. G-7 fiscal restructuring in the 1990s: macroeconomic effects. *Economic Policy* 20, 111- 146.

[7] Bayoumi, Tamim, Coe, D.T., and Helpman, E., 1996. R&D Spillovers and Global Growth. *IMF Working Paper* No. WP/96/47, International Monetary Fund, Washington.

[8] Becker, R, Rustem, B, 1993. Algorithms for solving nonlinear dynamic decision models. *Annals of Operations Research* 44, 117- 142.

[9] Black, R., Laxton, D., Rose, D., and Tetlow, R., 1994. The steady state model: SSQPM. The Bank of Canada's New Quarterly Projection Model, Part 1, Technical Report No. 72, Ottawa, Canada.

[10] Blanchard, O.J., and Kahn, C.M., 1980. The solution of linear difference models under rational expectations. *Econometrica* 48, 1305- 1311.

[11] Boucekkine, R., 1995. An Alternative Methodology for Solving Nonlinear Forward-Looking Models. *Journal of Economic Dynamics and Control* 19 (4), 711-734.

[12] Spectral??

[13] Boyd(1982a, J Comp Phys)

[14] Boyd(1982a, b, "Stieltjes" Math Comp)

[15] P. R. Spalart(1984, Contemp. Math.)

[16] Brown, P.N, and Saad, Y, 1990. Hybrid Krylov methods for nonlinear systems of equations. *SIAM Journal for Scientific and Statistical Computing* 11 (3), 450-481.

[17] Bruaset, A.M., 1996. Efficient solution of linear equations arising in a nonlinear economic model. In: Gilli, M. (Ed.), Computational Economic Systems. Models, Methods & Econometrics. Series: *Advances in Computational Economics*. Kluwer Academic Publishers, Dordrecht, pp. 243- 255.

[18] Bryant, R.C., Hooper, P., and Mann, C.L., 1993. *Evaluating policy regimes: New research in emperical macroeconomics*, The Brookings Institution, Washington.

[19] Coletti, D., Hunt, B., Rose, D., and Tetlow, R., 1996. The dynamic model: QPM. The Bank of CanadaÕs New Quarterly Projection Model, Part 3, Technical Report No. 75, Ottawa, Canada.

[20] Concus, P., Golub, G., and Meurant, G., 1985. Block preconditioning for the conjugate gradient method. *SIAM Journal for Scientific and Statistical Computing* 6, 220- 252.

[21] Dembo, R.S, Eisenstat, S.C, and Steihaug, T, 1982. Inexact Newton methods. *SIAM Journal for Numerical Analysis* 19 (2), 400- 408.

[22] Dennis, J.E. Jr., and Schnabel, R.B., 1983. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Series in Computational Mathematics*. Prentice-Hall, Englewood Cliffs, NJ.

[23] Don, F.J.H., and Gallo, G.M., 1987. Solving large sparse systems of equations in econometric models. *Journal of Forecasting* 6, 167- 180.

[24] Edison, H.J., Marquez, J.R., and Tryon, R.W., 1987. The structures and properties of the federal reserve board multicountry model. *Economic Modeling* 4, 115-315.

[25] Fair, R.C., and Taylor, J.B., 1983. Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models. *Econometrica* 51 (4), 1169-1185.

[26] Fisher, P.G., 1990. *SLIM: User's Manual For Ver. 81*, University of Warwick.

[27] Fisher, P.G., 1992. *Rational Expectations in Macroeconomic Models*. Kluwer Academic Publishers, Dordrecht.

[28] Fisher, P.G., Holly, S., and Hughes Hallett, A.J., 1986. Efficient solution techniques for dynamic non-linear rational expectations models. *Journal of Economic Dynamics and Control* 10, 139- 145.

[29] Fisher, P.G., and Hughes Hallett, A.J., 1987. The convergence characteristics of iterative techniques for solving econometric models. *Oxford Bulletin of Economics and Statistics* 49, 231- 244.

[30] Fisher, P.G., and Hughes Hallett, A.J., 1988a. Iterative techniques for solving simultaneous equation systems. a view from the economics literature. *Journal of Computational and Applied Mathematics* 24, 241- 255.

[31] Freund, R.W, 1993. A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems. *SIAM Journal on Scientific Computing* 14, 470- 482.

[32] Freund, R.W., Golub, G.H., and Nachtigal, N.M., 1991. Iterative solution of linear systems. *Acta Numerica*, 1- 44.

[33] Freund, R.W., and Nachtigal, N.M, 1991. QMR: a quasi-minimal residual method for non-Hermitian linear systems. *Numerische Mathematik* 60, 315- 339.

[34] Gagnon, J.E., 1991. A Forward-Looking Multicountry Model for Policy Analysis: MX3. *Economic and Financial Computing* 1, 311- 361.

[35] Gilbert, J.R, Moler, C, and Schreiber, R, 1992. Sparse matrices in MATLAB: design and implementation. *SIAM Journal on Matrix Analysis and Applications* 13, 333- 356.

[36] Gilli, M., 1992. Causal ordering and beyond. *International Economic Review* 33, 957- 971.

[37] Gilli, M, and Pauletto, G, 1997. Sparse direct methods for model simulation. *Journal of Economic Dynamics & Control* 21, 1093- 1111.

[38] Gilli, M., and Pauletto, G., 1998. Nonstationary iterative methods for solving models with forward looking variables. *Journal of Economic Dynamics and Control*, this issue.

[39] Hall, S.G., 1985. On the solution of large economic models with consistent expectations. *Bulletin of Economic Research* 37, 157- 161.

[40] Hall, S.G., 1987. Forecasting with a large macro model incorporating pervasive consistent expectations: NIESR Model 8. *Large Scale Systems Theory and Applications 13*, 145- 155.

[41] Helliwell, J.F., Meredith, G., Bagnoli, P., and Durand, Y., 1990. INTERMOD 1.1: A G-7 Version of the IMF's Multimod. *Economic Modelling* 7, 3- 62.

[42] Hestenes, M.R, and Stiefel, E, 1952. Method of conjugate gradients for solving linear systems. *Journal of Research of National Bureau Standards* 49, 409- 436.

[43] Hollinger, P., 1996. The Stacked-Time Simulator in TROLL: A robust algorithm for solving forward-looking models. Paper presented at the 2nd Int. Conf. on Computing in Economics and Finance, Geneva, Switzerland.

[44] Hughes Hallett, A.J., 1981. Some extensions and comparisons in the theory of Gauss- Seidel iterative techniques for solving large equation systems. In: Charatsis, E.G. (Ed.), Proceedings of the 1979 Econometric Society Meeting. North-Holland, Amsterdam.

[45] Hughes Hallett, A.J., 1984. Simple and optimal extrapolations for first order iterations. *International Journal of Computer Mathematics* 15, 309- 318.

[46] Hughes Hallett, A.J., 1986. The convergence of accelerated overrelaxation iterations. *Mathematics of Computation* 47, 219- 223.

[47] Hughes Hallett, A.J., and Fisher, P.G., 1990. On economic structures and model solution methods: Or should econometricians use newton methods for model solution? *Oxford Bulletin of Economics and Statistics* 52 (3) 317- 330.

[48] Hughes Hallett, A.J., Ma, Y., and Yin, Y.P., 1996. Hybrid algorithms with automatic switching for solving nonlinear equations systems in economics. *Journal of Economic Dynamics and Control* 20, 1051-1071.

[49] Hughes Hallett, A., and Piscitelli, L., 1998. A new convergence theorem for successive overrelaxation iterations, *Computational Economics* (forthcoming).

[50] Hughes Hallett, A.J., and Piscitelli, L., 1998. Simple reordering techniques for expanding the convergence radius of first order iterative techniques. *Journal of Economic Dynamics & Control* (this issue).

[51] Juillard, M., 1996. DYNARE: A program for the resolution and simulation of dynamic models with forward variables through the use of a relaxation algorithm. *CEPREMAP Working Paper* No. 9602, Paris, France.

[52] Juillard, M., Laxton, D., McAdam, P., and Pioro, H., 1998. An algorithm competition: first-order iterations versus Newton-based techniques. *Journal of Economic Dynamics & Control* (this issue).

[53] Kelley, C.T, 1995. *Iterative Methods for Linear and Nonlinear Equations.* SIAM, Philadelphia.

[54] Laffargue, J.P., 1990. Resolution d'un mode'le macroeconomique avec anticipations rationnelles. *Annales d'Economie et Statistique* 17, 97- 119.

[55] Laxton, D., Meredith, G., and Rose, D., 1995. Asymmetric effects of economic activity on inflation. *IMF Staff papers* Vol. 42, No. 2 International Monetary Fund, Washington, 344- 374.

[56] Luenberger, D.G., 1989. *Linear and Nonlinear Programming.* Addison-Wesley, Reading, MA.

[57] Masson, P.R., Symansky, S., and Meredith, G., 1990. MULTIMOD MARK II: A revised and extended model. *IMF Occasional Paper* No. 71, International Monetary Fund, Washington.

[58] McKibbin, W.J., and Sachs, J.D., 1991 Global linkages: Macroeconomic interdependance and cooperation in the world economy, The Brookings Institution, Washington.

[59] Paige, C, and Saunders, M, 1975. Solution of sparse indefinite systems of linear equations. *SIAM Journal of Numerical Analysis* 12, 617- 629.

[60] Pauletto, G., 1995. Solution and simulation of macroeconometric models, Thesis, Department of Econometrics, University of Geneva, Switzerland.

[61] Pauletto, G, 1997. *Computational Solution of Large-Scale Macroeconometric Models. Series: Advances in Computational Economics.* Kluwer Academic Publishers, Dordrecht.

[62] Pioro, H., McAdam, P., and Laxton, D., 1996. Solving MULTIMOD with first-Order and Newton-Based Techniques. University of Strathclyde Discussion Paper Working.

[63] Press, W., Teukolsky, S., Vetterling, W., and Flannery, B., 1992. Numerical Recipes In C: The Art of Scientific Computing. 2nd ed., Cambridge University Press, New York.

[64] Reid, J.K., 1977. Sparse Matrices. In: Jacobs, D. (Ed.), *The State of the Art in Numerical Analysis.* London Academic Press, , pp. 85- 146.

[65] Saad, Y., 1996. Iterative Methods for Sparse Linear Systems. PWS Publishing Comp., Boston, MA.9

[66] Saad, Y, and Schultz, M, 1986. GMRES: a generalized minimal residual algorithm for solving nonsymetric linear systems. *SIAM Journal for Scientific and Statistical Computing* 7, 856- 869.

[67] Steward, D.V., 1962. On an approach to techniques for the analysis of the structure of large systems. *SIAM Review* 4, 321- 342.

[68] Taylor, J.B., 1988. *The Treatment of Expectations in Large Multicountry Econometric Models.* Bryant, Henderson, and others, 161- 182.

[69] van der Vorst, H.A, 1992. Bi-CGSTAB: a fast and smoothly converging variant of BI-CG for the solution of nonsymetric linear systems. *SIAM Journal for Scientific and Statistical Computing* 13 (2) 631- 644.

[70] Young, D.M., 1971. *Iterative Solution of Large Linear Systems.* Academic Press, New York.