COMPUTATIONAL TECHNIQUES FOR ECONOMETRICS AND ECONOMIC ANALYSIS

Computational Techniques for Econometrics and Economic Analysis

edited by

D. A. Belsley

Boston College, Chestnut Hill, U.S.A.

KLUWER ACADEMIC PUBLISHERS

DORDRECHT / BOSTON / LONDON

Table of Contents

Preface vii

Part One: The Computer and Econometric Methods

Computational Aspects of Nonparametric Simulation Estimation
Ravi Bansal, A. Ronald Gallant, Robert Hussey, and George Tauchen 3

On the Accuracy and Efficiency Of GMM Estimators: A Monte Carlo Study
A. J. Hughes Hallett and Yue Ma 23

A Bootstrap Estimator for Dynamic Optimization Models
Albert J. Reed and Charles Hallahan 45

Computation of Optimum Control Functions by Lagrange Multipliers
Gregory C. Chow 65

Part Two: The Computer and Economic Analysis

Computational Approaches to Learning with Control Theory
David Kendrick 75

Computability, Complexity and Economics
Alfred Lorn Norman 89

Robust Min–Max Decisions with Rival Models
Berç Rustem 109

Part Three: Computational Techniques for Econometrics

Wavelets in Macroeconomics: An Introduction
William L. Goffe 137

MatClass: A Matrix Class for C++
C. R. Birchenhall 151

Parallel Implementations of Primal and Dual Algorithms for Matrix Balancing
Ismail Chabini, Omar Drissi-Kaïtouni and Michael Florian 173

RAVI BANSAL, A. RONALD GALLANT, ROBERT HUSSEY AND GEORGE TAUCHEN

Computational Aspects of Nonparametric Simulation Estimation

ABSTRACT. This paper develops a nonparametric estimator for structural equilibrium models that combines numerical solution techniques for nonlinear rational expectations models with nonparametric statistical techniques for characterizing the dynamic properties of time series data. The estimator uses the the score function from a nonparametric estimate of the law of motion of the observed data to define a GMM criterion function. In effect, it forces the economic model to generate simulated data so as to match a nonparametric estimate of the conditional density of the observed data. It differs from other simulated method of moments estimators in using the nonparametric density estimate, thereby allowing the data to dictate what features of the data are important for the structural model to match. The components of the scoring function characterize important kinds of nonlinearity in the data, including properties such as nonnormality and stochastic volatility.

The nonparametric density estimate is obtained using the Gallant-Tauchen seminonparametric (SNP) model. The simulated data that solve the economic model are obtained using Marcet's method of parameterized expectations. The paper gives a detailed description of the method of parameterized expectations applied to an equilibrium monetary model. It shows that the choice of the specification of the Euler equations and the manner of testing convergence have large effects on the rate of convergence of the solution procedure. It also reviews several optimization algorithms for minimizing the GMM objective function. The Nelder-Mead simplex method is found to be far more successful than others for our estimation problem.

1. INTRODUCTION

A structural equilibrium model is a complete description of a model economy including the economic environment, the optimization problem facing each agent, the market clearing conditions, and an assumption of rational expectations. A structural equilibrium model is difficult to estimate, as doing so entails repeated solution of a fixed-point problem in many variables. One approach is to employ a linearization, typically linear-quadratic, in conjunction with Gaussian specification for the errors. A linear specification is attractive because a closed form solution can be obtained (Hansen and Sargent, 1980). However, recent advances in numerical techniques now make it possible to obtain good approximate solutions for nonlinear models. (See the 1990 symposium in the *Journal of Business and Economic Statistics* (*JBES*), summarized in Tauchen, 1990 and Taylor and Uhlig, 1990.) At the same time as these developments in structural modelling have occurred, purely statistical models, such as ARCH (Engle, 1982), GARCH (Bollerslev, 1986), and seminonparametric

D. A. Belsley (ed.), Computational Techniques for Econometrics and Economic Analysis, 3–22.

models (Gallant and Tauchen, 1989, 1992), have been used to discover and characterize important forms of nonlinear behavior in economic time series, especially in financial time series. Linear Gaussian models cannot explain such nonlinear behavior in actual data. Thus, nonlinear structural models must be examined to see the extent to which they can explain the nonlinear behavior found in actual economic data. This paper shows how statistical techniques can be combined with numerical solution techniques to estimate nonlinear structural equilibrium models.

The most common approach for estimation of nonlinear structural models is probably generalized method of moments (GMM) applied to Euler equations, as developed in Hansen and Singleton (1982). This technique has been widely employed in financial economics and macroeconomics, though it is a limited information method and has shortcomings. For example, the estimation can encounter problems when there are unobserved variables, as is the case for the model we consider in Section 2 where the decision interval is a week, but some of the data are observed monthly. Also it does not provide an estimate of the law of motion of the economic variables. Thus, if the model is rejected, little information is available regarding the properties of the observed data that the model has failed to capture.

In this paper we describe an alternative strategy for estimating nonlinear structural models that was first applied in Bansal, Gallant, Hussey, and Tauchen (1992). The approach is similar to the simulated method of moments estimators of Duffie and Singleton (1989) and Ingram and Lee (1991). However, unlike those estimators, which match preselected moments of the data, our estimator minimizes a GMM criterion based on the score function of a nonparametric estimator of the conditional density of the observed data. In effect, the estimator uses as a standard of comparison a nonparametric estimate of the law of motion of the observed data. By selecting the GMM criterion in this way, we allow the observed data to determine the dynamic properties the structural model must match.

The estimator works by combining the method of parameterized expectations for numerically solving a nonlinear structural equilibrium model (Marcet, 1991; den Haan and Marcet, 1990) with the seminonparametric (SNP) method for estimating the conditional density of actual data (Gallant and Tauchen, 1989, 1992). For a particular setting of the parameters of the structural model, the method of parameterized expectations generates simulated data that solve the model. The model parameters are then estimated by searching for the parameter values that minimize a GMM criterion function based on the scoring function of the SNP conditional density estimate. The nonparametric structural estimator thus has three components: (1) using SNP to estimate the conditional density of actual data, (2) using the method of parameterized expectations to obtain simulated data that satisfy the structural model, and (3) estimating the underlying structural parameters by using an optimization algorithm that finds those parameter values that minimize the GMM criterion function.

Below we discuss in detail how the estimator works in the context of a two-country equilibrium monetary model. The model is based on Lucas (1982), Svensson (1985), and Bansal (1990), and is developed in full detail in Bansal, Gallant, Hussey, and Tauchen (1992). It accommodates time non-separabilities in preferences (Dunn and Singleton, 1986) and money via a transactions cost technology (Feenstra, 1986). In

effect, the model is a nonlinear filter that maps exogenous endowment and money supply processes into endogenous nominal processes, including exchange rates, interest rates, and forward rates. We show how this nonlinear dynamic model can be solved and simulated for estimation and evaluation.

In applying our estimator to this model, we find that there are several choices available to the researcher that greatly affect the estimator's success and rate of convergence. For example, the form in which one specifies the Euler equations on which the parameterized expectations algorithm operates can significantly affect the speed of convergence. This is an important finding, since our estimator uses this algorithm repeatedly at different model parameter values. Also, the means for testing convergence can have important consequences; we find it best to test for convergence of the projection used in parameterized expectations instead of testing for convergence of the coefficients representing the projection. Finally, we find that the complexity of our estimation procedure causes some optimization algorithms to have greater success in minimizing the GMM objective function. Among the optimization techniques we tried are gradient search methods, simulated annealing, and simplex methods. In Section 3.1 below we discuss how these methods work and their strengths and weaknesses for our type of optimization problem.

The rest of the paper is organized as follows: Section 2 specifies the illustrative monetary model and describes the simulation estimator. Section 3 discusses practical aspects of implementing the estimator, including solving the model with parameterized expectations and optimizing the GMM objective function to estimate the model parameters. Concluding remarks comprise the final section.

2. THE NONPARAMETRIC STRUCTURAL ESTIMATOR

2.1. *The Structural Model*

We apply our nonparametric structural estimator to the equilibrium monetary model of Bansal, Gallant, Hussey, and Tauchen (1992). In that model, a representative world consumer has preferences defined over services from two consumption goods. The utility function is assumed to have the form

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\left(c_{1t}^{*\delta} c_{2t}^{*1-\delta} \right)^{1-\gamma} - 1 \right] / (1-\gamma),$$

where $0 < \beta < 1$, $0 < \delta < 1$, $\gamma > 0$, and where c_{1t}^* and c_{2t}^* are the consumption services from goods produced in countries 1 and 2, respectively. Preferences are of the constant relative risk aversion (CRRA) type in terms of the composite consumption goods. The parameter γ is the coefficient of relative risk aversion, δ determines the allocation of expenditure between the two services, and β is the subjective discount factor. If $\gamma = 1$, then preferences collapse to log-utility

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\delta \ln c_{1t}^* + (1-\delta) \ln c_{2t}^* \right).$$

The transformation of goods to services is a linear technology

$$c_{1t}^* = c_{1t} + \kappa_{11}c_{1,t-1} + .. + \kappa_{1L_c}c_{1,t-L_c},$$

$$c_{2t}^* = c_{2t} + \kappa_{21}c_{2,t-1} + .. + \kappa_{2L_c}c_{2,t-L_c},$$

where c_{1t} and c_{2t} are the acquisitions of goods, the κ_{ij} determine the extent to which past acquisitions of goods provide services (and hence utility) in the current period, and L_c is the lag length. If $L_c = 0$, then the utility function collapses to the standard time separable case where $c_{1t}^* = c_{1t}$ and $c_{2t}^* = c_{2t}$. If the nonseparability parameters κ_{ij} are positive, then past acquisitions of goods provide services today. If they are negative, then there is habit persistence. Other patterns are possible as well. Recent acquisitions of goods can provide services today, while acquisitions further in the past contribute to habit persistence.

We introduce money into the model via a transaction-costs technology. The underlying justification for transactions costs is that the acquisition of goods is costly both in terms of resources and time. Money, by its presence, economizes on these costs and hence is valued in equilibrium. Transaction costs, $\psi(c, m)$, in our model are an increasing function of the amount of goods consumed c and a decreasing function of the magnitude of real balances m held by the consumer in the trading period. The functional form we use for the transaction-costs technology is

$$\psi(c, m) = \psi_0 c^{\alpha} m^{1-\alpha},$$

where $\psi_0 > 0$ and $\alpha > 1$.

The consumer's problem is to maximize expected utility $E_0\Sigma_{t=0}^{\infty}\beta^t U(c_{1t}^*, c_{2t}^*)$ by choosing $c_{1t}, c_{2t}, M_{1,t+1}, M_{2,t+1}, b_{1,t+1}^k$, and $b_{2,t+1}^k$, $k = 1, \ldots, N_a$, at time t subject to a sequence of budget constraints

$$\begin{aligned} &P_{1t}[c_{1t} + \psi(c_{1t}, m_{1t})] + e_t P_{2t}[c_{2t} + \psi(c_{2t}, m_{2t})] \\ &\quad + \sum_{k=1}^{N_a}(1/R_{1t}^k)b_{1,t+1}^k + \sum_{k=1}^{N_a}(f_t^k/R_{1t}^k)b_{2,t+1}^k + M_{1,t+1} + e_t M_{2,t+1} \\ &\leq \sum_{k=1}^{N_a}(1/R_{1t}^{k-1})b_{1t}^k + \sum_{k=1}^{N_a}(f_t^{k-1}/R_{1t}^{k-1})b_{2t}^k + M_{1t} + e_t M_{2t} \\ &\quad + P_{1t}w_{1t} + e_t P_{2t}w_{2t} + q_{1t} + e_t q_{2t}. \end{aligned}$$

Here, P_{1t} and P_{2t} are current prices of consumption goods c_{1t} and c_{2t} in the units of the respective country's currency. $M_{1,t+1}$ and $M_{2,t+1}$ are the stocks of currency in the two countries carried forward from period t to $t + 1$. Real money balances, $m_{1t} = M_{1t}/P_{1t}$ and $m_{2t} = M_{2t}/P_{1t}$, are defined in terms of beginning of period money holdings. The $b_{1,t+1}^k$ and $b_{2,t+1}^k$ are the agent's holdings of risk-free claims to the currencies of countries 1 and 2 in period $t + k$. Claims on country 1's currency are made by trading pure discount bonds with gross k-period interest rates R_{1t}^k. Claims on country 2's currency are made by trading forward contracts in the currency market,

where e_t is the spot exchange rate and f_t^k is the k-period forward exchange rate, with both rates defined in units of country 1's currency per unit of country 2's currency. w_{1t} and w_{2t} are the stochastic endowments of goods within the two countries. Lump sum transfers of q_{1t} and q_{2t} units of currency are made by the government at time t. These transfers are known to the agent at the beginning of period t but can be used for carrying out transactions only in period $t+1$.

The stationary decision problem facing the agent delivers the following Euler equations for the asset holdings $M_{1,t+1}$ and $M_{2,t+1}$:

$$E_t\Big[MU_{c_{it}} - \beta MU_{c_{i,t+1}}\Big(\frac{P_{it}}{P_{i,t+1}}\Big)\Big(\frac{1+\psi_{c_{it}}}{1+\psi_{c_{i,t+1}}}\Big)\Big(1-\psi_{m_{i,t+1}}\Big)\Big] = 0, \quad i = 1,2,$$

and for $b_{1,t+1}^k$ and $b_{2,t+1}^k$:

$$E_t\Big[MU_{c_{1t}} - \beta^k MU_{c_{1,t+k}}\Big(\frac{P_{1t}}{P_{1,t+k}}\Big)\Big(\frac{1+\psi_{c_{1t}}}{1+\psi_{c_{1,t+k}}}\Big)R_{1t}^k\Big] = 0,$$

$$E_t\Big[MU_{c_{2t}} - \beta^k MU_{c_{2,t+k}}\Big(\frac{P_{2t}}{P_{2,t+k}}\Big)\Big(\frac{1+\psi_{c_{1t}}}{1+\psi_{c_{2,t+k}}}\Big)\Big(e_t R_{1t}^k / f_t^k\Big)\Big] = 0,$$

where $MU_{c_{it}}$ is the marginal utility of c_{it}, and $\psi_{c_{it}}$ and $\psi_{m_{it}}$ are the derivatives of transaction costs, $\psi(c_{it}, m_{it})$, with respect to the first and second arguments, respectively. Transactions costs modify the returns to the two monies, M_{1t} and M_{2t}. We would expect $P_{1t}/P_{1,t+1}$ to be the return at time $t+1$ for carrying forward an extra unit of country one's currency today. However, because of transaction costs, every extra unit of currency carried forward also lowers transaction costs in the next period by a real amount, $-\psi_{m_{i,t+1}}$, so the total return is given by $[(1-\psi_{m_{i,t+1}})P_{it}/P_{i,t+1}]$. The model also delivers an intratemporal restriction on the choice of goods c_{1t} and c_{2t}

$$e_t = E_t\Big[\Big(\frac{MU_{c_{2t}}}{MU_{c_{1t}}}\Big)\Big(\frac{P_{1t}}{P_{2t}}\Big)\Big(\frac{1+\psi_{c_{1t}}}{1+\psi_{c_{2t}}}\Big)\Big].$$

In maximizing utility, the consumer faces an exogenous stochastic process that governs the evolution of money growth and endowment growth in the two countries. We define the operator d to produce the ratio of the value of a variable in one period to its value in the previous period, as, for example, $dM_{1t} = M_{1t}/M_{1,t-1}$. Using this operator, we specify a driving process for the exogenous state vector $S_t = (dM_{1t}, dM_{2t}, dw_{1t}, dw_{2t})$ of the form

$$\log S_t = a_0 + A \log S_{t-1} + u_t,$$

where u_t is $iid\ N(0, \Omega)$, a_0 is a 4-vector, and A and Ω are 4×4 matrices. More complex stochastic processes for the exogenous state variables could easily be accommodated by our numerical solution method.

The final elements needed to complete the description of the model are the market clearing conditions

$$c_{it} + \psi(c_{it}, m_{it}) = w_{it},$$

$$M_{i,t+1} = q_{it} + M_{it}\,, \qquad i = 1, 2.$$

The parameter vector of the structural economic model is

$$\lambda = (\beta, \gamma, \delta, \psi_0, \alpha, \kappa_{11}, \ldots, \kappa_{1L_c}, \kappa_{21}, \ldots, \kappa_{2L_c}, a_0', \; vec(A)', \; vech(\Omega^{1/2})')'.$$

For each value of λ the model defines a nonlinear mapping from the strictly exogenous process $\{S_t\}$ to an output process $\{U_t\}$. The output process is

$$U_t = (dM_{1t}, dM_{2t}, dw_{1t}, dw_{2t}, dc_{1t}, dc_{2t}, dP_{1t}, dP_{2t}, R^4_{1t}, f^4_t/e_t, de_t)',$$

which is an 11-vector containing the elements of S_t along with the gross consumption growth rates, the gross inflation rates, the four-period interest rate in country 1, the ratio of the four-period forward exchange rate to the spot rate, and the gross growth rate of the spot exchange rate. It proves convenient also to include the elements of S_t in the output process, mapping them directly with an identity map. The particular set of variables comprising the remaining elements of U_t are those endogenous variables that turn out to be of interest for various aspects of the analysis of the model and the empirical work.

The mapping from $(\{S_t\}, \lambda)$ to the endogenous elements of U_t is defined by the solution to the nonlinear rational expectations model. In practice, we use Marcet's method of parameterized expectations (Marcet, 1991; den Haan and Marcet, 1990) to approximate the map. Given a value of λ, the method "solves" the model in the sense of determining simulated realizations of the variables that satisfy the Euler equations. In what follows, $\{U_t^\lambda\}$ denotes a realization of the output process given λ and a realization of $\{S_t\}$. A complete description of how we apply the method of parameterized expectations to this problem is given in Section 3.1 below.

2.2. *The Estimation Method*

The nonlinearity of the economic model prevents estimation by traditional methods since it is computationally intractable to compute the likelihood of a sample as a function of the model's parameters. However, simulation methods can be used to compute predicted probabilities and expectations under the model. Thus we propose a new simulation estimator that estimates the model by searching for the value of the parameter λ for which the dynamic properties of data simulated from the model match, as closely as possible, the properites of actual data.

Not all elements of U_t generated by the model are actaully observed weekly, so our empirical strategy is to use latent-variable methods with our simulation estimator. High quality observations on financial market prices, i.e., payoff data, are widely available on a weekly basis, and so we concentrate on these series in the estimation. We utilize weekly observations on three raw series: SPOT_t, the spot exchange rate (in \$ per DM); FORWARD^4_t, the 30-day forward rate (in \$ per DM); and TBILL^4_t, the one month treasury bill interest rate, computed from the term structure, and quoted on a bank discount basis. From the raw series we form a 3-element process $y_t = (y_{1t}, y_{2t}, y_{3t})'$ with

$$y_{1t} = 100 * \log(\text{SPOT}_t/\text{SPOT}_{t-1}),$$

$$y_{2t} = 100 * \log(\text{FORWARD}_t^4/\text{SPOT}_t),$$
$$y_{3t} = \text{TBILL}_t^4.$$

Exploratory empirical work indicates that $\{y_t\}$ is reasonably taken as a strictly stationary process, while the levels of the exchange rate series are nonstationary.

The correspondence between the elements of y_t and those of the output vector U_t are as follows: Country 1 is the U.S. and country 2 is Germany. Given a simulated realization $\{U_t^\lambda\}$ from the model, the corresponding $\{y_t^\lambda\}$ is computed as

$$y_{1t}^\lambda = 100 * \log(de_t^\lambda),$$
$$y_{2t}^\lambda = 100 * \log(f_t^4/e_t),$$
$$y_{3t}^\lambda = 100 * (360/30)[1 - (1/R_{1t}^{4\lambda})].$$

The expression for y_{3t}^λ converts $1/R_{1t}^{4\lambda}$, which is the price at time t of \$1 in period $t+4$, to an annualized interest rate using the bank discount formula customarily applied to treasury bill prices (Stigum, 1990, p. 66).

The observed process is $\{y_t\}$ and the simulated process is $\{y_t^\lambda\}$ as defined above. The $\{y_t\}$ process is computed directly from the raw data while $\{y_t^\lambda\}$ is computed using the structural model of Section 2.1. We assume the model to be "true" in the sense that there is a particular value, λ_0, of the structural parameter vector and a realization, $\{S_{0t}\}$, of the exogenous vector such that the observed $\{y_t\}$ is obtained from $(\{S_{0t}\}, \lambda_0)$ in exactly the same manner that the model generates $\{y_t^\lambda\}$ from $(\{S_t\}, \lambda)$.

In broad terms, the estimation problem of this paper is analogous to the situation described, among others, by Duffie and Singleton (1989) and Ingram and Lee (1991). Common practice in such situations is to use a simulated method of moments estimator of λ_0 based on certain *a priori* selected moments of the data. We likewise propose such an estimator, but we take a different approach in determining what moments to match and in assigning relative weights in matching those moments.

The estimation strategy of this paper starts from the point of view that the structural model should be forced to confront all empirically relevant aspects of the observed process. The observed process $\{y_t\}$ is strictly stationary and possibly nonlinear, so its dynamics are completely described by the one-step ahead conditional density $f(y_t|\{y_{t-j}\}_{j=1}^\infty)$. Let $\hat{f}(\cdot|\cdot)$ denote a consistent nonparametric estimate computed from a realization $\{y_t\}_{t=t_0}^n$. The estimator $\hat{f}(\cdot|\cdot)$ defines what is empirically relevant about the process and thereby provides a comprehensive standard of reference upon which to match the economic model to the data.

The keystone to our structural estimator is the scoring function of the SNP estimator of Gallant and Tauchen (1989, 1992), which provides a consistent nonparametric estimator of the conditional density under mild regularity conditions. This use of the nonparametric fit to define the criterion of estimation motivates our choice of the term "nonparametric structural estimator". The Gallant-Tauchen estimator is a truncation estimator based on a series expansion that defines an hierarchy of increasingly complex models. The estimator $\hat{f}(\cdot|\cdot) = f_K(\cdot|\cdot, \hat{\theta}_{K_n})$ is characterized by an auxiliary parameter vector $\hat{\theta}_{K_n}$ that contains the coefficients of the expansion; the

subscript K denotes the K^{th} model in the hierarchy. The length of $\hat{\theta}_{K_n}$ depends on the model. In practice, K is determined by a model selection criterion that slowly expands the model with sample size n and thereby ensures consistency. For the K^{th} model in the hierarchy, the corresponding $\hat{\theta}_{K_n}$ solves the first-order condition

$$\frac{\partial}{\partial \theta_{K_n}} \mathcal{L}_{K_n}(\{y_t\}_{t=t_0}^{n}, \hat{\theta}_{K_n}) = 0,$$

where $\mathcal{L}_{K_n}(\cdot)$ is the sample log likelihood of the corresponding model.

The nonparametric structural estimator is defined by mimicking this condition. Specifically, subject to identifiability conditions, a consistent estimator is available by choosing $\hat{\lambda}$ to make the same condition hold (as closely as possible) in the simulation

$$\frac{\partial}{\partial \theta_{K_n}} \mathcal{L}_{K_n}(\{y_\tau^\lambda\}_{\tau=\tau_0}^{T}, \hat{\theta}_{K_n}) \approx 0.$$

The left-hand side is the gradient of the log likelihood function evaluated at a simulated realization $\{y_\tau^\lambda\}_{\tau=\tau_0}^{T}$ and at the $\hat{\theta}_{K_n}$ determined by fitting the K^{th} SNP model to the actual data $\{y_t\}_{t=t_0}^{n}$. If the length of λ, ℓ_λ, is less than the length of θ_K, ℓ_K, then the model is overidentified (under the order condition) and a GMM criterion is used to minimize the length of the left-hand side with respect to a suitable weighting matrix.

Interestingly, this approach defines a consistent and asymptotically normal estimator irrespective of the particular SNP model used, so long as $\ell_K \geq \ell_\lambda$ and an identification condition is met. In practice, we implement the estimator using the particular SNP model that emerges from the specification search in the nonparametric estimation of $f(\cdot|\cdot)$. The choice of K is thus data-determined. This selection rule forces the scoring function to be appropriate for the particular sample at hand. The scoring function of the fitted SNP model contains just those indicators important to fit the data and no more. Also, because the fitted SNP model has the interpretation of a nonparametric maximum-likelihood estimator, the information equality from maximum likelihood theory provides a convenient simplification that greatly facilitates estimation of the weighting matrix for the GMM estimation.

3. IMPLEMENTING THE ESTIMATOR

In this section we discuss the practical aspects of implementing the nonparametric structural estimator described above. The implementation entails an initial SNP estimation of the conditional density of observed payoff data. The score function from this density estimate defines what properties our nonparametric structural estimator must mimic. Because estimating SNP models has been described extensively in Gallant and Tauchen (1989, 1992), we do not review that procedure here.

Following the SNP estimation, there are three distinct components to the procedure. The first involves using the method of parameterized expectations to solve the structural model for a particular value of the parameter vector λ. The second entails combining the initial SNP estimation with the parameterized expectations procedure

to form the GMM objective function for the nonparametric structural estimator. The third is optimization of the objective function. Each of these components is described in detail below.

3.1. *Solving the Model Using Parameterized Expectations*

We use the method of parameterized expectations (Marcet, 1991; den Haan and Marcet, 1990) to obtain simulated data that satisfy the Euler equations of the structural economic model. In essence, this method approximates conditional expectations of certain terms with the projections of those terms on a polynomial in the state variables. The method uses Euler equations to iterate between postulated values of time series and projections based on those postulated values until those values and projections each converge. This procedure will be explained more fully below.

We find that the specification of the Euler equations greatly affects the speed with which the parameterized expectations algorithm converges. From Section 2, the first two Euler equations are

$$E_t\left[MU_{c_{it}} - \beta MU_{c_{i,t+1}}\left(\frac{P_{it}}{P_{i,t+1}}\right)\left(\frac{1+\psi_{c_{it}}}{1+\psi_{c_{i,t+1}}}\right)\left(1-\psi_{m_{i,t+1}}\right)\right] = 0, \quad i = 1,2.$$

Using the definition of the velocity of money, $V_{it} = c_{it}P_{it}/M_{it}$, $i = 1,2$, one form in which these equations can be rewritten is

$$V_{it} = \frac{E_t\left(MU_{c_{it}}\right)}{E_t\left[\beta MU_{c_{i,t+1}}\left(\frac{dc_{i,t+1}}{dM_{i,t+1}V_{i,t+1}}\right)\left(\frac{1+\psi_{c_{it}}}{1+\psi_{c_{i,t+1}}}\right)\left(1-\psi_{m_{i,t+1}}\right)\right]}, \quad i = 1,2.$$

Because of the time nonseparabilities in our model, it is also possible to rearrange these Euler equations into an alternative form that expresses velocity as a single conditional expectation rather than the ratio of two conditional expectations. (We omit the derivation here.) It would seem at first that expressing the Euler equations as a single conditional expectation would be advantageous since the solution algorithm would have to estimate only one conditional expectation per Euler equation rather than two. However, we have found that convergence of the algorithm with this specification is much slower. This occurs because the single conditional expectation contains a difference of two terms that remains stable across iterations, while the time series from which it is constructed moves around substantially. The conditional expectation of this difference is less informative for updating guesses at the solution time series than are the two conditional expectations specified in the ratio above.

The next step in setting up the Euler equations entails various mathematical manipulations that allow them to be expressed in terms of conditional expectations of functions of velocity, consumption growth, and money growth:

$$V_{it} = \frac{E_t\left[f_{i1}\left(dc_{1,t-L_c+1}, dc_{2,t-L_c+1}, \ldots, dc_{1,t+L_c}, dc_{2,t+L_c}; \lambda\right)\right]}{E_t\left[f_{i2}\left(dc_{1,t-L_c+2}, dc_{2,t-L_c+2}, \ldots, dc_{1,t+L_c+1}, dc_{2,t+L_c+1}, V_{it}, V_{i,t+1}, dM_{i,t+1}; \lambda\right)\right]}$$

$$= \frac{E_t(F_{i1,t})}{E_t(F_{i2,t})} \qquad i = 1, 2,$$

where the $f_{ij}(\cdot)$ are particular functional forms too complex to be written out here. The market clearing conditions of the model imply that

$$dc_{it} = g(V_{it}, V_{i,t-1}, dw_{it}; \lambda) = dw_{it}(1 + \psi_0 V_{i,t-1}^{\alpha-1})/(1 + \psi_0 V_{it}^{\alpha-1}), \quad i = 1, 2$$

Given a vector λ and a realization of the exogenous state variables S_t – which includes money growth dM_{it} and endowment growth dw_{it} – consumption growth dc_{it} is an exact function of velocity, so the Euler equations above are fixed-point equations in the two velocity series. This means we can solve these first two Euler equations for the two equilibrium velocity processes as a unit before considering the remaining Euler equations. Using the solution velocity processes, we can then calculate directly equilibrium consumption growth dc_{it} and inflation dP_{it} for the two countries, and we can solve the remaining Euler equations to determine the equilibrium k-period interest rates in country 1, R_{1t}^k, the premium of the k-period forward rate over the spot rate f_t^k/e_t and exchange rate growth de_t.

Several methods have been used to solve nonlinear rational expectations models with endogenous state variables (Taylor and Uhlig, 1990; Judd, 1991). Among these, parameterized expectations is particularly suited to use with a simulation estimator because it produces simulated data that satisfy the Euler equations without having to solve for the full decision rule. We parameterize each of the conditional expectations in the above Euler equations as a function of the exogenous and endogenous state variables. The augmented vector of state variables is

$$\tilde{S}_t = (1, S_t, dc_{1,t-1}, dc_{2,t-1}, \cdots, dc_{1,t-L_c+1}, dc_{2,t-L_c+1}),$$

where 1 is concatenated for use as a constant in the regressions described below. If $L_c \leq 1$, then there are no endogenous state variables, and $\tilde{S}_t$ is just equal to S_t and a constant. Any class of dense functions, such as polynomials or neural nets, can be used to approximate the conditional expectations. The particular functional form we use to parameterize expectations is

$$E_t(F_{ij,t}) = \exp[\text{poly}(\tilde{S}_t, \nu_{ij})],$$

where poly$(\cdot)$ is a polynomial in $\tilde{S}_t$, and ν_{ij} is the vector of its coefficients. We choose to use an exponential polynomial because economic theory implies that $E_t[F_{ij,t}]$ should be positive. In practice, the polynomial we use consists of linear and squared terms of the elements of $\tilde{S}_t$.

Below is a description of the algorithm for solving for the equilibrium velocity series given a vector λ. In every instance, the ranges of the indices are $i = 1, 2$ and $j = 1, 2$; superscripts indicate iteration numbers.

Step 1.
Simulate a realization of $\{u_t\}$, where u_t is $iid\ N(0, \Omega)$.

Step 2.
From some initial S_0, generate a realization of $\{S_t\}$ using

$$\log S_t = a_0 + A \log S_{t-1} + u_t.$$

In practice, we set S_0 to a vector of ones, but in performing the parameterized expectations regressions we exclude the first five hundred observations from the simulated data to eliminate any effect from choosing initial values.

Step 3.
Determine starting realizations of the velocity series $\{V_{it}^0\}$.

We consider two possible ways to do this. The first is to specify starting values for ν_{ij}^0, perhaps values of ν_{ij} obtained from a previous solution at a nearby λ. Then, given ν_{ij}^0 and some initial observations on velocities $V_{it}^0, t = 0, \ldots, L_c$, the remaining elements of the starting velocity series for $t = L_c + 1, \ldots, T$, can be determined using the following relationships recursively

$$dc_{it}^0 = g(V_{it}^0, V_{i,t-1}^0, dw_{it}; \lambda),$$

$$V_{it}^0 = \exp[\text{poly}(\tilde{S}_t^0, \nu_{i1}^0)] / \exp[\text{poly}(\tilde{S}_t^0, \nu_{i2}^0)].$$

This structure is recursive because $\tilde{S}_t^0$ contains $dc_{i,t-1}^0$. A drawback to this approach is that the simulated time series produced by the solution procedure are dependent upon the starting values, so any attempt to replicate the solution exactly would require knowing those starting values.

A second approach for establishing starting realizations of the velocity series would set V_{1t}^0 and V_{2t}^0 to be constants for all t. For these constants, one could calculate steady-state values for the two velocities, or simply set the velocities equal to 1. This latter approach still produces convergence in a relatively small number of iterations.

Regardless of the approach used to determine starting values of velocity, if one uses the procedure described below to improve the stability of the algorithm by dampening iteration updates, starting values must also be specified for the polynomial coefficients ν_{ij}^0. We recommend setting all of the coefficients to zero except the constants. This means that $E_t(F_{ij,t}) = \exp[\text{poly}(\tilde{S}_t, \nu_{ij})]$ reduces to $E_t(F_{ij,t}) = \exp[\text{constant}_{ij}]$. The constants can be set equal to the log of the unconditional means of the $F_{ij,t}$'s. Setting the initial polynomial coefficients in this way gives a very stable position from which to start the iterations.

Step 4.
Iteration k: Using the V_{it}^{k-1} series, calculate the $F_{ij,t}^{k-1}$ and regress each of these four on a linearized version of $\exp[\text{poly}(S_t^{k-1}, \nu_{ij}^k)]$ to estimate ν_{ij}^k. The linearization is done around ν_{ij}^{k-1}.

A linearized version of the exponential function is used to allow one to perform linear regressions rather than nonlinear regressions at each iteration. When the

coefficients converge ($\nu_{ij}^k = \nu_{ij}^{k-1}$), the value of the exponential function is equal to the value of its linearized version at the point at which we want to evaluate it.

Den Haan and Marcet (1990) actually suggest a more gradual way of modifying the guesses at the polynomial coefficients from iteration to iteration. Rather than setting ν_{ij}^k equal to the coefficients obtained from the regressions, one can set ν_{ij}^k equal to a convex combination of those coefficients, call them b_{ij}^k, and the guess at the coefficients from the previous iteration as

$$\nu_{ij}^k = \rho b_{ij}^k + (1-\rho)\nu_{ij}^{k-1},$$

where $0 < \rho \leq 1$. This procedure has the effect of dampening the speed with which the guesses at the coefficients are updated. The smaller is ρ, the more gradually the coefficients are modified from one iteration to the next. One might want to use this gradual updating scheme to stabilize iterations that are not well behaved. For the model in this paper, we were always able to set $\rho = 1$, which implies no dampening in updating the coefficients.

Step 5.
Determine the two V_{it}^k series according to

$$V_{it}^k = \exp[\text{poly}(\tilde{S}_t^{k-1}, \nu_{i1}^k)] / \exp(\text{poly}[\tilde{S}_t^{k-1}, \nu_{i2}^k)],$$

and the two dc_{it}^k series according to

$$dc_{it}^k = g(V_{it}^k, V_{i,t-1}^k, dw_{it}; \lambda).$$

Step 6.
Repeat steps 3 and 4 until the velocity series converge. Convergence is reached when

$$\max_i \max_t |(V_{it}^k - V_{it}^{k-1})/(V_{it}^{k-1} + \epsilon)| \leq \xi,$$

where ϵ and ξ are small positive numbers.

Note that we check convergence on the velocity series, that is, on the ratios of the parameterized expectations projections, which is a different procedure than that used in Marcet (1991). Marcet looks for convergence of the coefficients of the projections, rather than of the projections themselves. We check convergence on the projections because of complications that arise when there is a high degree of multicollinearity between the variables of the parameterized polynomial, as is the case in our model. Multicollinearity makes it possible for the coefficients of the polynomial to continue to oscillate between successive iterations even though the projection onto the polynomial has essentially converged. Since it is the values of the projections that are important for solving the model, we look for convergence of those values.

In summary, the parameterized expectations solution method works by alternating between estimating values of conditional expectations based on some postulated

realization of the velocity processes (which amounts to estimating the ν_{ij}'s) and updating the postulated values of the velocity processes based on the estimated conditional expectation values. The procedure continues until the velocity processes converge.

Once the equilibrium velocity and consumption growth series have been determined from the first two Euler equations, the four-period interest rate series in country 1, the premium of the four-period forward exchange rate over the spot rate, and the exchange rate growth can be determined from the remaining Euler equations without additional iterations. The Euler equations can be written as

$$R_{1t}^4 = \frac{E_t(F_{11,t})/(1+\alpha\psi_0 V_{1t}^{\alpha-1})}{\beta^4 E_t\left[f_{13}\left(dc_{1,t+1}, dc_{2,t+1}, \ldots, dc_{1,t+L_c+5}, dc_{2,t+L_c+5}, dP_{1,t+1}, \ldots, dP_{1,t+4}; \lambda\right)\right]},$$

$$e_t R_{1t}^4 / f_t^4 = \frac{E_t(F_{21,t})/(1+\alpha\psi_0 V_{2t}^{\alpha-1})}{\beta^4 E_t\left[f_{23}\left(dc_{1,t+1}, dc_{2,t+1}, \ldots, dc_{1,t+L_c+5}, dc_{2,t+L_c+5}, dP_{2,t+1}, \ldots, dP_{2,t+4}; \lambda\right)\right]},$$

$$de_t = \frac{E_t(F_{21,t})E_{t-1}(F_{11,t-1})(1+\alpha\psi_0 V_{1t}^{\alpha-1})(1+\alpha\psi_0 V_{2,t-1}^{\alpha-1})dP_{1t}dc_{1t}}{E_{t-1}(F_{21,t-1})E_t(F_{11,t})(1+\alpha\psi_0 V_{2t}^{\alpha-1})(1+\alpha\psi_0 V_{1,t1}^{\alpha-1})dP_{2t}dc_{2t}}.$$

In these equations $dP_{it} = (dM_{it}V_{it})/(dc_{it}V_{i,t-1})$, the gross inflation rate in each country. As before, f_{13} and f_{23} are particular function forms. The conditional expectations terms in the equations are each estimated by regressing the value of the function inside the expectations operator on a polynomial in $\tilde{S}_t$. The polynomial we use consists of the elements of $\tilde{S}_t$ raised to the first, second, third, and fourth powers. The resulting simulation values are used to form $\{y_t^\lambda\}$.

The time required to solve the structural economic model at some value of λ is an important consideration, since our nonparametric estimator requires solutions at many different values of λ in finding the value that minimizes the GMM objective function. When we use simulated time series of length 1000 to solve the model (excluding an initial discarded 500 observations), convergence for most values of λ is achieved in approximately one minute on a SUN SPARCstation 2.

3.2. *Defining the GMM Objective Function*

The Gallant-Tauchen (1992) SNP estimator underlies our nonparametric structural estimator. Following their notation, given the observed process $\{y_t\}$, let $x_{t-1} = (y'_{t-1}, \ldots, y'_{t-L})'$ and let $p(y_t|x_{t-1}, \lambda_0)$ denote the conditional density of y_t conditional on L lags of itself and the true λ_0. By stationarity, we can suppress the t subscript and simply write $p(y|x, \lambda_0)$ when convenient. In addition, let $p(y, x, \lambda_0)$ denote the joint density of (y_t, x_{t-1}). Frequently, we suppress the dependence of

the conditional density on λ_0 and write $p(y|x)$, but we always make explicit the dependence of the joint density $p(y, x, \lambda_0)$ on λ_0, because that becomes important.

The SNP estimator is a sieve estimator that is based on the sequence of models $\{f_K(y|x, \theta_K)\}_{K=0}^{\infty}$, where $\theta_K \in \Theta_K \subseteq \Re^{t_K}$, $\Theta_K \subseteq \Theta_{K+1}$ and where $f(y|x, \theta_K)$ is a truncated Hermite series expansion. This hierarchy of models can, under regularity conditions, approximate $p(y|x)$ well in the sense

$$\lim_{K\to\infty} \inf_{\theta_K \in \Theta_K} \|[f_K(\cdot|\cdot, \theta_K) - p(\cdot|\cdot)\| = 0,$$

where $\|\cdot\|$ is a Sobelov norm. The approximation also holds along a sequence of estimated models fitted to data sets $\{y_{-L+1}, \ldots, y_n\}$, $n = 1, 2, \ldots \infty$, with the the appropriate model for each n determined by a model selection strategy.

The key component of our nonparametric structural estimator is the mean gradient of the log-density of a K^{th} order SNP model,

$$g(\lambda, \theta_K) = \int\int (\partial/\partial\theta) \log[f_K(y|x, \theta_K)]p(y, x, \lambda)dydx.$$

In practice, the above expectation is approximated by simulating $\{U_\tau^\lambda\}_{\tau=1}^{T}$, forming $\{y_\tau^\lambda\}$ as just described, taking lags to form $\{x_{\tau-1}^\lambda\}$, and then averaging

$$\tilde{g}(\lambda, \theta_K) = \frac{1}{T}\sum_{\tau=1}^{T}(\partial/\partial\theta) \log[f_K(y_\tau^\lambda|x_{\tau-1}^\lambda, \theta_K)].$$

We take $\tilde{g}(\lambda, \theta_K) \approx g(\lambda, \theta_K)$.

The nonparametric structural estimator is defined as follows: Let $\{y_t\}_{t=-L+1}^{n}$ be a realization of the observed process and let

$$\hat{\theta}_{Kn} = \arg\max_{\theta_K \in \Theta_K} \left\{\frac{1}{n}\sum_{t=1}^{n} \log[f(y_t|x_{t-1}, \theta_K)]\right\}.$$

Thus, $\hat{\theta}_{Kn}$ is the estimated parameter vector of a K^{th} order SNP model fitted to the data by maximum likelihood. The estimator $\hat{\lambda}$ is the solution of the GMM estimation problem

$$\hat{\lambda} = \arg\min_{\lambda \in \Lambda}\{\tilde{s}_n(\lambda)\},$$

where

$$\tilde{s}_n(\lambda) = n\tilde{g}(\lambda, \hat{\theta}_{Kn})' W_n \tilde{g}(\lambda, \hat{\theta}_{Kn}),$$

and where W_n is a symmetric positive definite weighting matrix such that $W_n \to W$ almost surely and W is positive definite. In the application, we use

$$W_n = \left\{\frac{1}{n}\sum_{t=1}^{n}(\partial/\partial\theta) \log[f(y_t|x_{t-1}, \hat{\theta}_{Kn})](\partial/\partial\theta') \log[f(y_t|x_{t-1}, \hat{\theta}_{Kn}])\right\}^{-1}.$$

which is the natural estimate of the inverse of the information matrix based on the gradient-outer-product formula. This choice makes the minimized value of the GMM objective function, $\tilde{s}_n(\lambda)$, approximately $\mathcal{X}^2(\ell_K - \ell_\lambda)$ for large K.

Below we consider several different algorithms for minimizing $\tilde{s}_n(\lambda)$. Regardless of the algorithm, it is advantageous to control the interface between the optimizer and the economic model by scaling the optimizer's guesses at the parameter values to be within a range in accordance with the economic theory behind our model. For example, in our model it only makes sense for δ to be between 0 and 1, so we constrain the optimizer to attempt solutions only with such values. These constraints are imposed by using various forms of logistic transformations.

3.3. *Optimizing the Objective Function*

The basic computational task for the estimator is to evaluate $\hat{\lambda} = \text{argmin}_{\lambda \in \Lambda} \{\tilde{s}_n(\lambda)\}$. This minimization is not straightforward for our problem because of the large number of parameters to be estimated (between 37 and 41 depending upon whether one, two, or three lags of consumption services enter the utility function) and because analytical derivatives of the objective function with respect to λ are not available. We tried four different algorithms for minimizing the objective function and found significant differences across algorithms for our problem.

3.3.1. *Optimizing with NPSOL and DFP*

We initially tried two classic gradient search methods: NPSOL (Gill, Murray, Saunders, and Wright, 1986), and Davidon-Fletcher-Powell (DFP), as implemented in the GQOPT package (Quandt and Goldfeld, 1991). Both algorithms work in a similar manner. A search direction is determined, a one-dimensional optimization is performed along that direction, and then the search direction is updated. The process is repeated until a putative optimum is achieved.

These algorithms work quite well when analytic derivatives are available. For example, we use NPSOL to perform the preliminary SNP parameter estimation to compute $\hat{\theta}_{Kn}$, which is needed to form $\tilde{s}_n(\lambda)$. Analytical derivatives are available for the SNP objective function, and NPSOL works adequately even on fairly large problems. In our application, the SNP estimation itself entails a specification search over roughly thirty different models with some having as many as 150 parameters. That whole effort takes only three or four days on a SUN SPARCstation 2. In an variety of other SNP applications, NPSOL has been found to work reasonably well (Gallant and Tauchen, 1992).

Analytical derivatives of $\tilde{s}_n(\lambda)$, however, are computationally infeasible. The process $\{y_t^\lambda\}$ is a solution to a fixed-point problem, as are its analytical derivatives. Computing $\partial \tilde{s}_n(\lambda)/\partial \lambda$ would involve computing a solution to a fixed-point problem for each component. Evaluating $\tilde{s}_n(\lambda)$ and its derivatives for arbitrary λ is well beyond the reach of current computing equipment. The large computational demands for analytical derivatives appear to be intrinsic to all solution methods for nonlinear structural models, including those described in the *JBES* Symposium (Tauchen, 1990; Taylor and Uhlig, 1990) or Judd (1991), since they all entail solving nonlinear fixed-

point problems.

Gradient search methods use numerical derivatives in place of analytical derivatives when the latter are unavailable. For our type of problem, this does not work well. The computations turn out to be about as demanding as would be those for analytical derivatives approximating the gradient of the objective function $\partial \tilde{s}_n(\lambda)/\partial\lambda$ at a single point λ entails computing the simulated process $\{y_t^\lambda\}$ after small perturbations in each of the λ. With ℓ_λ on the order of 37 to 41, this entails, at a minimum, recomputing the equilibrium of the model that many additional times just to approximate a single one-sided gradient. The net effect is to generate about as many function calls as would a naive grid-search. In fact, our experience suggests that a naive grid search might even work better. In the course of approximating $\partial \tilde{s}_n(\lambda)/\partial\lambda$ via perturbing λ and forming difference quotients, values of λ that produce sharp improvement in the objective function are uncovered quite by happenstance. Neither NPSOL nor DFP retains and makes use subsequently of these particularly promising values of λ; the effort that goes into to computing the equilibrium for these λ is lost. Simple grid search would retain these λ's.

3.3.2. *Optimizing with Simulated Annealing*

We also tried simulated annealing, a global method. An implementation of simulated annealing by William Goffe is available in the GQOPT optimization package (Quandt and Goldfeld, 1991). We used an updated version that William Goffe kindly made available to us. See Goffe, Ferrier, and Rodgers (1992) for a discussion of the algorithm and additional references. We give a brief summary of the essential ideas here.

From a point λ, simulated annealing changes element i of λ using

$$\lambda_i' = \lambda_i + r v_i,$$

where r is a uniformly distributed random number over $[-1, 1]$ and v_i is the ith element of a vector of weights V. If $\tilde{s}_n(\lambda')$ is smaller than $\tilde{s}_n(\lambda)$ the point is accepted. If not, the point is accepted if a random draw from the uniform over $[0, 1]$ exceeds

$$p = e^{[\tilde{s}_n(\lambda') - \tilde{s}_n(\lambda)]/T}.$$

The elements of V and T are tuning parameters that must be selected in advance and are adjusted throughout the course of the iterations. We used the defaults. There are additional tuning parameters that determine when these adjustments occur. Again, we accepted the defaults.

The algorithm was defeated by the large number of function evaluations that it requires. Most exasperating was its insistence on exploring unprofitable parameter values. After making some promising initial progress the algorithm would plateau far from an optimum and give no indication that further progress could be achieved if the iterations were permitted to continue.

3.3.3. *Optimizing with Simplex Methods*

The optimization method that performs best for our problem is the simplex method developed by Nelder and Mead (1964). Fortran code for implementing this method

is available in the GQOPT optimization package (Quandt and Goldfeld, 1991). The method works as follows: We begin the minimization of a function of ℓ_λ variables by constructing a simplex of $(\ell_\lambda + 1)$ points in ℓ_λ-dimensional space: $\lambda_0, \lambda_1, \ldots, \lambda_{\ell_\lambda}$. We denote the value of the function at point λ_i by s_i. The lowest, highest, and second highest values are

$$s_l = \min_i(s_i), \qquad s_h = \max_i(s_i), \qquad s_{hh} = \max_{i \neq h}(s_i),$$

corresponding to points λ_l, λ_h, and λ_{hh}. We also define the notation $[\lambda_i \lambda_j]$ to indicate the distance from λ_i to λ_j.

The algorithm works by replacing λ_h in the simplex continuously by another point with a lower function value. Three operations are used to search for such a new point–*reflection*, *contraction*, and *expansion*–each of which is undertaken relative to the centroid $\bar{\lambda}$ of the simplex points excluding λ_h. The centroid is constructed as

$$\bar{\lambda} = \frac{1}{\ell_\lambda} \sum_{i=1}^{\ell_\lambda} \lambda_i, \qquad i \neq h.$$

The *reflection* of λ_h through the centroid is λ_r, which is defined by

$$\lambda_r = (1 + \alpha_r)\bar{\lambda} - \alpha_r \lambda_h,$$

where $\alpha_r > 0$ is the *reflection coefficient*. λ_r lies on the line between λ_h and $\bar{\lambda}$, on the far side of $\bar{\lambda}$, and α_r is the ratio of the distance $[\lambda_r \bar{\lambda}]$ to $[\lambda_h \bar{\lambda}]$. If $s_l < s_r \leq s_{hh}$, we replace λ_h with λ_r and start the process again with this new simplex.

If reflection has produced a new minimum ($s_r < s_l$), we search for an even lower function value by expanding the reflection. The *expansion* point is defined by

$$\lambda_e = \alpha_e \lambda_r + (1 - \alpha_e)\bar{\lambda}$$

where $\alpha_e > 1$ is the *expansion coefficient* that defines the ratio of the distance $[\lambda_e \bar{\lambda}]$ to $[\lambda_r \bar{\lambda}]$. λ_e is farther out than λ_r on the line between λ_h and $\bar{\lambda}$. If $s_e < s_r$, λ_h is replaced in the simplex by λ_e. Otherwise, the expansion has failed and λ_r replaces λ_h. The process is then restarted with the new simplex.

If reflection of λ_h has not even produced a function value less than s_{hh} – which means that replacing λ_h with λ_r would leave s_r the maximum – we rename λ_h to be either the old λ_h or λ_r, whichever has a lower function value. Then we attempt to find an improved point by constructing the *contraction*

$$\lambda_c = \alpha_c \lambda_h + (1 - \alpha_c)\bar{\lambda},$$

where $0 < \alpha_c < 1$. The *contraction coefficient* α_c is the ratio of the distance $[\lambda_c \bar{\lambda}]$ to $[\lambda_h \bar{\lambda}]$. If $s_c < s_h$, then the contraction has succeeded, and we replace λ_h with λ_c and restart the process. If this contraction has failed, we construct a new simplex by contracting all the points toward the one with the lowest function value, which is accomplished by replacing the λ_i's with $(\lambda_i + \lambda_l)/2$. Then the process of updating the simplex restarts.

Nelder and Mead suggest stopping their procedure when the standard deviation of the λ_i's is less than some critical value. In our empirical work, we strengthen this stopping rule by restarting the algorithm several times from the value on which the Nelder-Mead procedure settles. When this restarting leads to no further significant improvement in the objective function value, we accept the best point as the minimum of the function. In implementing the algorithm, we also found it advantageous to modify the error handling procedures of the Nelder-Mead code provided in GQOPT slightly to allow us to start the procedure with a wider ranging simplex.

The Nelder-Mead simplex method was far more successful than the other methods we tried for minimizing our objective function. There are two aspects of this method that we believe are responsible for its success. First, the method finds new lower points on the objective surface without estimating derivatives. Second, by using the operations of *reflection*, *expansion*, and *contraction*, the Nelder-Mead method is designed to jump over ridges in the objective surface easily in searching for new lower points. This property can be important in preventing an optimization algorithm from shutting down too early. Despite these advantages, however, the performance of the Nelder-Mead method is not completely satisfactory, because it requires a very large number of function calls to find the minimum of the function. Given the number of parameters in our model and the complexity of evaluating the objective function at any one point, the method can occupy several weeks of computing time on a Sun SPARCstation. Even though this computing demand is substantial and far greater than we expected from the outset of this project, we still consider our nonparametric structural estimator very successful in achieving our goal of estimating a nonlinear rational expectations model and fully accounting for the complex nonlinear dynamics of actual time series in that estimation.

Results from applying this estimator to the illustrative monetary model are available in Bansal, Gallant, Hussey, and Tauchen (1992).

4. CONCLUSION

In this paper we describe a new nonparametric estimator for structural equilibrium models and show its application to an equilibrium monetary model. The discussion of the implementation of the estimator indicates important considerations that might arise in applying the estimator to other nonlinear rational expectations models.

There are several advantages to this estimator. By using the method of parameterized expectations to solve the model numerically, structural equilibrium models can be estimated without limiting oneself to linear approximations. By using a consistent nonparametric estimate of the conditional density of the observed data to define the criterion to be minimized in estimation, the estimator forces the model to confront the law of motion of the observed data, which can include complex forms of nonlinearity. Finally, the estimator provides simulated data from the model. If a model is rejected, then it is possible to evaluate the dimensions in which it fails to match characteristics of the observed data, thus providing valuable diagnostic information for building better models.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grants No. SES-8808015 and SES-90-23083. We thank Geert Bekaert, Lars Hansen, David Hsieh, Ellen McGrattan, Tom Sargent, and many seminar and conference participants for helpful comments at various stages of this research.

REFERENCES

Bansal, R., 1990, "Can non-separabilities explain exchange rate movements and risk premia?", Carnegie Mellon University, Ph.D. dissertation.

Bansal, R., A. R. Gallant, R. Hussey and G. Tauchen, 1992, "Nonparametric estimation of structural models for high-frequency currency market data", Duke University, manuscript.

Bollerslev, T., 1986, "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics* 31, 307–327.

den Haan, W. J. and A. Marcet, 1990, "Solving the stochastic growth model by parameterizing expectations", *Journal of Business and Economic Statistics* 8, 31–4.

Duffie, D. and K. J. Singleton, 1989, "Simulated moments estimation of markov models of asset prices", Stanford University, Graduate School of Business, manuscript.

Dunn, Kenneth and K. J. Singleton, 1986, "Modeling the term structure of interest rates under non-separable utility and durability of goods", *Journal of Financial Economics* 17, 27–55.

Engle, R. F., 1982, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica* 50, 987–1007.

Feenstra, R. C., 1986, "Functional equivalence between liquidity costs and the utility of money", *Journal of Monetary Economics* 17, 271–291.

Gallant, A. R. and G. Tauchen, 1989, "Seminonparametric estimation of conditionally constrained heterogeneous processes: asset pricing applications", *Econometrica* 57, 1091–1120.

Gallant, A. R. and G. Tauchen, 1992, "A nonparametric approach to nonlinear time series: estimation and simulation", in David Brillinger, Peter Caines, John Zeweke, Emanuel Paryen, Murray Rosenblatt, and Murad S. Taggu (eds.), *New Directions in Time Series Analysis, Part II*, New York: Springer-Verlag, 71–92.

Gill, P. E., W. Murray, M. A. Saunders and M. H. Wright, 1986, "User's guide for NPSOL (version 4.0): a Fortran package for nonlinear programming", Technical Report SOL 86-2, Palo Alto: Systems Optimization Laboratory, Stanford University.

Goffe, W. L., G. D. Ferrier, and J. Rodgers, 1992, "Global Optimization of statistical functions: Preliminary results" in Hans M. Amman, David A. Belsley, and Louis F. Pau (eds.), *Computational Economics and Econometrics, Advanced Studies in Theoretical and Applied Econometrics*, Vol. 22, 19–32, Boston: Kluwer Academic Publishers.

Hansen, L. P., 1982, "Large sample properties of generalized method of moments estimators", *Econometrica* 50, 1029–1054.

Hansen, L. P. and T. J. Sargent, 1980, "Formulation and estimation of dynamic linear rational expectations models", *Journal of Economic Dynamics and Control* 2, 7–46.

Hansen, L. P. and K. J. Singleton, 1982, "Generalized instrumental variables estimators of nonlinear rational expectations models", *Econometrica* 50, 1269–1286.

Ingram, B. F. and B. S. Lee, 1991, "Simulation estimation of time-series models", *Journal of Econometrics* 47, 197–205.

Judd, K. L., 1991, "Minimum weighted least residual methods for solving aggregate growth models", Federal Reserve Bank of Minneapolis, Institute of Empirical Macroeconomics, manuscript.

Lucas, R. E., Jr., 1982, "Interest rates and currency prices in a two-country world", *Journal of Monetary Economics* 10, 335–360.
Marcet, A., 1991, "Solution of nonlinear models by parameterizing expectations: an application to asset pricing with production", manuscript.
McCallum, B. T., 1983, "On non-uniqueness in rational expectations models: an attempt at perspective", *Journal of Monetary Economics* 11, 139–168.
Nelder, J. A. and R. Mead, 1964, "A simplex method for function minimization", *The Computer Journal* 7, 308–313.
Quandt, R. E. and S. M. Goldfeld, 1991, *GQOPT/PC*, Princeton, N.J.
Stigum, M., 1990, *The money market*, 3rd ed., Homewood, Il.: Dow Jones–Irwin.
Svensson, L. E. O., 1985, "Currency prices, terms of trade and interest rates: a general equilibrium asset-pricing cash-in-advance approach", *Journal of International Economics* 18, 17–41.
Tauchen, G., 1990, "Associate editor's introduction", *Journal of Business and Economic Statistics* 8, 1.
Taylor, J. B. and H. Uhlig, 1990, "Solving nonlinear stochastic growth models: a comparison of alternative solution methods", *Journal of Business and Economic Statistics* 8, 1–17.