Discussion Paper 15

Institute for Empirical Macroeconomics
Federal Reserve Bank of Minneapolis
250 Marquette Avenue
Minneapolis, Minnesota 55480

May 1989

SOLVING NONLINEAR STOCHASTIC OPTIMIZATION AND EQUILIBRIUM PROBLEMS BACKWARDS

Christopher A. Sims
Institute for Empirical
Macroeconomics and
University of Minnesota

ABSTRACT

In a stochastic equilibrium model some stochastic processes are usually exogenously given, while other are either chosen optimally by agents or emerge from market equilibrium conditions. When we simulate such a model, often we aim at studying the relations among variables in the model as we vary parameters of policy and of behavior of economic agents. We are no more certain (indeed often less certain) of what is reasonable or interesting behavior for the exogenous variables (some of which may be unobservable) than of the variables chosen by agents or fixed in markets. It turns out that if we are flexible about which variables' behavior we take as given in the model solution computation, freeing ourselves from the convention that the variables exogenous to the model economy must be taken as given in the simulation computations, great computational savings may result.

This material is based on work supported by the National Science Foundation under Grant No. SES-8722451. The Government has certain rights to this material.

Any opinions, findings, conclusions, or recommendations expressed herein are those of the author(s) and not necessarily those of the National Science Foundation, the University of Minnesota, the Federal Reserve Bank of Minneapolis, or the Federal Reserve System.

SOLVING NONLINEAR STOCHASTIC OPTIMIZATION AND EQUILIBRIUM PROBLEMS BACKWARDS

by Christopher A. Sims

May 1989

Economists have in recent years been applying the assumption that economic agents dynamically optimize under uncertainty. Models embodying this assumption can be difficult to solve, however, and those which have been used have therefore been highly simplified or approximate. We have solution methods for linear-quadratic models, in which agents have linear decision rules, and for general nonlinear models with small discrete state spaces. Kydland and Prescott have shown the way toward approximating non-linear-quadratic models with linear-quadratic ones, thereby making them soluble, but their method is essentially equivalent to assuming that agents apply certainty equivalence to their problems. For some issues the effects of the failure of certainty equivalence are of central interest. This paper suggests a class of solution methods which allow analysis of a broader range of stochastic equilibrium models than has heretofore been practical.

Most of the methodology economists have applied in this area is borrowed from engineering. For an engineer, just as for an economic agent, solving the stochastic optimization problem requires a method for finding the appropriate decision rule by which to make controlled variables respond to a given forcing stochastic process. An economist, however, can take a more symmetric view of control variables and forcing variables. In modeling behavior, he is interested in the mapping between the stochastic processes for controlled and uncontrolled variables. Since he may observe data on both processes, both are equally "given" to him. In fact the variables taken as given by one group of agents in his model may well be taken as choice variables by

other agents. For an economist, therefore, a convenient method of generating solution pairs -- controlled with corresponding uncontrolled stochastic processes -- may be useful even if the method does not proceed by generating the controlled process corresponding to an arbitrary uncontrolled processes.

1. The Method in Linear-Quadratic Problems

We begin by discussing the application of these ideas to stationary linear-quadratic (LQ) models in state-space form. In this context backwards solution provides no computational advantage, but for economists used to LQ control problems, it is easiest to explain why the method works and what can go wrong with it in this elementary framework. And the method may have some value in theoretical and empirical exercises even here, if we have stronger a priori notions about the shape of a reasonable decision rule than about the distributional properties of unobservable disturbance terms.

We postulate an nx1 vector x(t) of variables to be chosen subject to the system of constraints

$$Ax(t+1) = Bx(t) + \varepsilon(t+1), \qquad (1)$$

with $E_t[\varepsilon(t+1)]=0$. If a dynamic system involves lags of higher than first order, it can be cast into the form (1) by the usual expansion of the x vector to include lagged values. The system (1) of course includes the canonical state-control framework in which the first part of the x vector is called the state vector and the A matrix has the form A=[I 0]. For our purposes the separation of states from controls is not essential, and we allow general A matrices. We follow the convention that variables

dated t are in the information set at t. We assume an objective function

$$E\left[\sum_{t=0}^{\infty}\beta^{t}x(t)'Mx(t)\right] \quad . \tag{2}$$

This implies first order conditions, in addition to (1), of

$$2Mx(t) + \beta B'E_{t}[\lambda(t+1)] = A'\lambda(t) \tag{3}$$

Introducing an error term $\eta(t)=\lambda(t)-E_{t-1}[\lambda(t)]$, we obtain

$$\begin{bmatrix} A & 0 \\ 0 & \beta B' \end{bmatrix}\begin{bmatrix} x(t+1) \\ \lambda(t+1) \end{bmatrix} = \begin{bmatrix} B & 0 \\ -2M & A' \end{bmatrix}\begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} + \begin{bmatrix} \varepsilon(t+1) \\ \beta B'\eta(t+1) \end{bmatrix} . \tag{4}$$

Letting $z = (x',\lambda')'$ and $\xi = (\varepsilon',\beta\eta'B)'$, (4) becomes

$$\Gamma_{0}z(t+1) = \Gamma_{1}z(t) + \xi(t+1) \tag{5}$$

Equation (5) is a stochastic difference equation system which will be satisfied by any solution to the LQ problem. However, it ordinarily cannot be used to generate simulated solutions from given initial conditions. The most immediate obstacle is that Γ_{0} is generally singular. Furthermore, stochastic processes z satisfying (5) generally will explode at a rate exceeding $\beta^{-t/2}$, which may make the objective function unbounded. If we impose the requirement that

$$E\left[\sum_{t=1}^{\infty}x(t)'x(t)\beta^{t}\right] < \infty \quad , \tag{6}$$

which guarantees boundedness of the objective function, we acquire as an additional first order condition the transversality condition

$$E\left[\sum \lambda(t)'\lambda(t)\beta^t \right] < \infty \ .^1 \qquad (7)$$

If there were no restrictions on the joint distribution of $\varepsilon(t)$ and $\eta(t)$, (5) would usually imply highly explosive paths for z, which would not be solutions to the optimization problem. Furthermore, since η is an artificial disturbance term, if there were no restrictions relating η to ε there would be nonuniqueness in the solution -- for a given ε we could generate a new solution for every possible choice of the η process satisfying $E_t[\eta(t+1)]=0$. The hard part of the solution is to find the functional relation between $\varepsilon(t)$ and $\eta(t)$ which makes solutions to (5) satisfy (6) and (7).

We can easily characterize the desired subset of solutions to (5) and obtain the decision rule for agents by using the QZ decomposition of (Γ_1,Γ_0), which is defined as follows. For any pair of square matrices (Γ_0,Γ_1), there are orthonormal matrices Q,Z and upper triangular matrices H_0,H_1 such that

$$\Gamma_0 = QH_0Z, \quad \Gamma_1 = QH_1Z. \qquad (8)$$

Furthermore we can always choose Q and Z such that all zeroes on the diagonal of H_0 occur in the lower right corner and such that the remaining ratios of diagonal elements H_{1ii}/H_{0ii} are monotone non-decreasing in absolute value as we proceed from upper left to

---

[1]The condition (6) is stronger than needed to guarantee boundedness of the objective function, since M need not be full rank. The problem may have a solution even with no side condition like (6) or with side conditions more general than (6), e.g. allowing different rates of growth for different components of x but we ignore this possibility here. A derivation of the first order conditions appears in an Appendix available from the author.

lower right down the diagonal. These ratios of diagonal elements are called the generalized eigenvalues of (Γ_1,Γ_0), and the first k rows of Z span the same space as the first k generalized right eigenvectors of the pair.[2]

We can use (8) to rewrite (5). We replace z(t) with $Z\gamma(t)$ and premultiply the system (5) by Q' to obtain

$$H_0\gamma(t+1) = H_1\gamma(t) + Q'\xi(t+1) \quad . \tag{9}$$

If there are zeroes on the diagonal of H_0 which correspond to zeroes on the diagonal of H_1, the system is ill-specified. Any solution to the system can be modified by adding to it arbitrary multiples of the corresponding generalized eigenvectors. Hence we will assume that this does not occur. Other zeroes on the diagonal of H_0 correspond to components of a solution path which the system suppresses. That is, at every date the solution must lie in the space spanned by the remaining columns of Z. This can be checked by observing that if there is a zero at the bottom of the diagonal of H_0, taking E_t of both sides of the last line of (9) gives us

---

[2] The program PC Matlab contains a command to compute the decomposition (7), except that it will not automatically order the eigenvalues along the diagonal. It does compute the generalized eigenvectors, which is all we need for this application if there are no repeated roots. Unfortunately 0's on the diagonals are likely to repeat in models of any size. Algorithms for rearranging eigenvalues along the diagonal of the Schur decomposition of a single matrix are displayed in Golub and Van Loan (1983), p. 240-245. These can be adapted to apply to (8), and a PC-Matlab routine which does this is available from the author.

$$0 = H_{1nn}\gamma_n(t) \quad . \tag{10}$$

But then, if there is a zero on the diagonal of H_0 in the n-1'st position, using (10) in the n-1'st line gives us $0=\gamma_{n-1}(t)$, etc. as long as we keep finding zeroes as we move up the diagonal of H_0.

When the last q diagonal elements of H_0, and hence the last q elements of γ, are zero, we have q redundant equations in (9). If p=n-q, then by construction the upper left pxp submatrix of H_0, which we will call G_0, is nonsingular, so we can drop the last q equations of (9) and retain a system capable of determining the p nonzero elements of $\gamma(t)$, which we will label $\tau(t)$. Using successive lagged versions of the top n-q rows of (9) to substitute for the lagged τ on the right-hand side, we can derive

$$\tau(t) = (G_0^{-1}G_1)^t\tau(0) + \sum_{s=0}^{t-1}(G_0^{-1}G_1)^sQ^{1\prime}\xi(t-s) \quad , \tag{11}$$

where Q^1 is the first n-q columns of Q.

From (11) it is evident that we can suppress all terms in τ, and hence in z, which grow at the rate $\beta^{-.5}$ or faster only by requiring that all components of the τ vector corresponding to generalized eigenvalues of (Γ_1,Γ_0) equal to or greater than $\beta^{-.5}$ in absolute value be zero. Because of the way we have ordered the matrices, if there are m eigenvalues greater than $\beta^{-.5}$, this amounts to setting the last m elements of τ to be identically zero, which in turn is just

$$Z_{j.}z(t) = 0, \text{ all } t, \; j=p-m+1,\ldots,p \quad , \tag{12}$$

where $Z_{j.}$ is the j'th row of Z. The conditions (12) implicitly impose also the requirement that the corresponding rows of $Q'\xi(t)$ be identically zero.

One way to simulate solution paths for the system, then, would be to use all but the last q (identically zero) rows of (11), together with (12). This system will have more equations than the length of the z vector, but we can use it by treating some elements of the ξ vector as unknowns. This is possible because the restrictions on ξ implied by (12) should make all the random variables η introduced into the first order conditions in arriving at (5) linear functions of the original disturbances ε. If they do not, the system allows multiple solutions for a given exogenous input. It is also possible for the implied restrictions on ξ to be inconsistent with the exogenously given properties of ε, but this implies that the problem as orginally posed does not have a solution. Assuming the system is well behaved in these respects, we can use the known properties of ε to generate realizations of the ε process, then use (11) and (12) to generate γ paths. Paths for z then emerge from $z=G\gamma$. The same procedure applies if instead of generating simulated z paths our problem is to control z in response to exogenously given ε's.

In economic applications we may sometimes have a clearer idea of what the stochastic properties of η should be than of what the stochastic properties of ε should be. We could, therefore, draw simulation realizations for η and solve the system for ε instead of drawing ε paths and solving for η. Of course since the mapping between ε and η is known and linear, which one we start with in simulation does not affect our ability to control the covariance matrix of either ε or η in the simulation.

The equations (12) define a mapping from some of the elements of z to others. Usually we think of this mapping as determining "control" variables from observations on "states" and we call the

mapping the "decision rule".[3] In some economic applications we might know something a priori about the form of the decision rule. Suppose we postulated a model (1)-(2), derived the decision rule mapping (12), and concluded that the result was unreasonable. What would we get if we simulated (5) with a different version of (12), based on our notion of a reasonable form for it?

If (5) and our revised (12) are used to solve for z and η from given ε, the result will be meaningless. If, however, we draw simulated paths for η which satisfy the martingale-difference property ($E_t[\eta_{t+1}]=0$), possibly also draw paths for some subsvector or linear transformation of ε, and use (5) and our revised (12) to solve for the remainder of ε, the result is still a solution to a model closely related to the one we started with. In particular, the original Euler equations and the constraints still hold, so the resulting simulation, if it satisfies the growth rate conditions (6) and (7), is a solution to a problem with the same objective function and same constraints. Since (12) is altered, something about the problem must have changed, though, and that can only be the distribution of ε, including its serial dependence properties. In this linear quadratic

---

[3] Actually, this is a slight distortion of the usual terminology. (12) relates elements of the z_t vector to one another. What is usually called a decision rule relates the elements of the z_t vector labeled "controls" to the elements of the z_t vector labeled "exogenous shocks" and to the elements of z_{t-1} labeled "states". That is, a conventional decision rule relates current controls to current shocks and lagged states, with the current state then determined from (1). In a conventional state space model, (1) can be solved for the current state and the result substituted in to (12) to yield the conventional decision rule.

framework, the only aspect of the distribution of ε which affects the solution is its serial correlation structure, so it is this which must be changed by the change in the decision rule.

The argument of the preceding paragraph makes an important general point which holds outside the linear-quadratic context -- if the Euler equations and the form of the constraints are held constant, changes in the stochastic properties of the disturbances in the constraints map into changes in the decision rule, and vice versa.

We may encounter two possible kinds of pathology in working backwards to properties of ε from the form of the decision rule. One obvious one is that it may turn out that when the system is solved for some or all of the ε's, taking η as input and using an arbitrary form for (12), the resulting path for z (and hence ε) may be explosive at a rate greater than $\beta^{-.5t}$. Since we have set the problem up so that we know that highly explosive paths for z are not solutions, we know that in this case we have generated a system in which, though the Euler equations are satisfied, we are not getting solutions to the optimum problem.

Another problem is that it may turn out that, though by construction $\varepsilon(t)$ is determined in the simulations by current and past values of those elements of the η and ε vectors being taken as input, it is not possible to represent $\eta(t)$ as determined by current and past values of $\varepsilon(t)$.[4] This does not mean that we are failing to generate optimal paths, but it does mean that the paths take as available at t a larger information set than the set of

---

[4]The simplest example of this phenomena is $\varepsilon(t)=\eta(t)-2\eta(t-1)$. It is easily verified that in this case $\eta(t)$ cannot be approximated arbitrarily well in mean square by linear combinations of $\varepsilon(s)$ for $s\leq t$.

current and past values of $\varepsilon(t)$. Putting it another way, if the model is taken as descriptive of some agents' behavior, those agents are being taken to be able to forecast $\varepsilon(t)$ more accurately than they could if they were able to observe past ε alone.

We call these two pathologies the stability and invertibility problems, respectively. They can be characterized as in a sense duals of each other. A stability problem arises when the equation system implies that the space spanned by current and past ε's under the covariance inner product is strictly larger than the space spanned by current and past values of the subset of (η,ε) taken as input. The invertibility problem arises when the relative size of the two spaces is reversed. When there is a stability problem the altered (12) used is not a valid decision rule for any version of ε's serial correlation properties, while the invertibility problem only implies that there may be some question about the solution's interpretation. The infinite horizon discounted LQ context may be a little misleading here, however. Even when the solution is stable and invertible, there is a question of whether the implied joint behavior of x and ε is reasonable based on introspection and observation of behavior.[5] When we turn below to examination of the finite horizon case, we will see that there is no longer any sharp distinction among

---

[5]This approach to the solution makes these issues come to the fore. But of course no matter how one solves the system, if it is a model of economic behavior one faces questions of whether the implied behavior is reasonable. Deriving properties of ε from properties of the decision rule makes clear in the numerical procedures a basic fact about the nature of economic science -- an economic model always includes arbitrary parameterizations which can be adjusted to make its behavior more or less reasonable.

stability problems, invertibility problems, and general problems of reasonableness of results.

II. Application to Non-LQ Problems

In non-LQ problems, the idea of backsolving solution is to use a decision rule -- an analogue to (12) -- together with the Euler equations and the constraints, to simulate a solution path. This is just the proposal for the LQ problem in the previous section. However in a non-LQ problem the distinction between generating paths for ε and solving for η on the one hand, and generating η paths and solving for ε on the other, is more important. Because in a non-LQ problem the mapping from ε to η is nonlinear and only implicit in the equation system, solving for ε from η may allow computationally cheap simulated paths for exact solutions, where direct solution is prohibitively expensive.

The reason non-LQ problems are so expensive to solve is that the analogue to (12) is ordinarily defined implicitly by functional equations which involve all aspects of the distribution of $\varepsilon(t)$. But if we guess a reasonable form for (12) and delete a corresponding number of equations from (1), the Euler equations (3) and the remaining constraints can be solved period-by-period to generate simulated solutions.

Of course we can still run in to stability and invertibility problems, but they can be minimized by centering the analysis on an approximating LQ problem. That is, we can begin by following Kydland and Prescott in expanding the equations of the problem in Taylor series about a steady state solution path, deriving the corresponding form for (12). Then we can replace the Euler equations and constraints of the LQ approximation with the nonlinear Euler equations and constraints of the original problem and append (12) (or, more usually, a convenient nonlinear equation

whose first-order expansion about the steady state solution matches (12)) to the resulting system. This system can be solved in the usual way, deleting some equations of the nonlinear analogue to (1) and deriving the simulated paths of ε from the assumed input paths for a subset of the (ε, η) vector. Since in the LQ problem itself we are sure that with (12) derived this way there is no stability or invertibility problem, we can be sure the same will hold true locally for the nonlinear problem, so long as stochastic disturbances are small enough and we have enough continuity.

The result of this approach to solution is a simulated path which is an approximate solution to the original non-LQ problem in the same sense that Kydland and Prescott's solution is an approximate solution. The solution is also an exact solution, though, to a problem with the same objective function and constraints, but a slightly different stochastic process for $\varepsilon(t)$ from that originally proposed. In situations where one is suspicious of the approximate certainty-equivalence assumption on which the Kydland-Prescott approximation is based, a comparison of the two types of solution may give important insights into the effects of failure of certainty-equivalence.

III. Example 1: One Sector Growth

We consider a simple growth model, cast as the maximization problem

$$\max E\left[\sum_{t=1}^{\infty} \frac{C(t)^{1-\gamma}}{1-\gamma} \beta^t \right] \tag{13}$$

subject to

$$C(t+1) + K(t+1) - K(t) = \theta(t) K(t)^{\alpha} \tag{14}$$

$$\log(\theta(t)) = \tau \log(\theta(t-1)) + \nu(t) \tag{15}$$

We assume $E_t[\nu(t+1)]=0$ as the starting point for our LQ approximation, but of course will not quite achieve this in our backsolved solution.

Euler equations are

$$C(t)^{-\gamma} = \lambda(t) \tag{16}$$

$$\lambda(t) = \beta E_t\left[\lambda(t+1)\left\{\alpha\theta(t+1)K(t)^{\alpha-1}+1\right\}\right] \tag{17}$$

This problem was solved and simulated for 1000 observations under two specifications[6]. In both, α=.33, β=.98 and θ was lognormally distributed, with τ=.95 in (15). In simulation I, Var(ν(t))=.0004 and γ=3. In simulation II, Var(ν(t))=.01 and γ=.5. Both the lower risk aversion and the higher variance of ε contribute to wider fluctuations in simulation II.

Each simulation was undertaken both with an LQ approximation around the steady state and with a backsolved simulation. The LQ approximation was constructed by solving the linearized system to arrive at (12), then simulating equations generated by taking the loglinear first order expansion of all equations of the system about the deterministic steady state. (12) itself, as well as the rest of the system, was used in loglinear form. If instead the system had been simulated as linear in levels, it would have under some parameter values and shock distributions implied negative consumption or capital.

---

[6]The nonlinear simulations are part of a set of 10 undertaken as a group project with members of the NBER-NSF seminar on computation for nonlinear rational expectations models. Members of the group attacked the same problem with a variety of solution methods. The results are submitted for publication.

The backsolved system, of course, used the Euler equation and constraints in the original nonlinear form. It did not use (12) directly, however. Because the system involves non-integer powers of K, it can become insoluble in the real domain if it generates negative K values. Direct use of (12) or a loglinear expansion of it is therefore capable of making the system insoluble. Instead the simulations described here replaced (12) with an equation loglinear in $C(t)/K(t)$, $K(t-1)$, and $\theta(t)$ with coefficients chosen to give the system the same Taylor expansion about steady state as would have been produced by direct use of (12).

The combination of high risk aversion and small shocks in simulation I results in narrow fluctuations around the steady state. Differences in the properties of the loglinear LQ approximation and the backsolved solution were small for this solution. This is not to say they were undetectable. Since the technology constraint (14) is not loglinear, it cannot hold exactly in the loglinear simulation. Chart 1.I shows a plot of "demand" $C(t)+K(t)-K(t-1)$ vs. "supply" $\theta(t)K(t-1)^{\alpha}$ for the simulation I loglinear LQ simulation. For a simulation which uses (14) in its nonlinear form, like the backsolved solution, the chart would have shown a perfect straight line. The LQ simulation plot has clearly visible, but fairly small, departures from a perfect straight line. Chart 1.II shows the same data for the loglinear LQ simulation II, and the deviations from the straight line are notably more pronounced.

Equations (16) and (17) together yield a stochastic equation in C, K, and θ, which implies that if we define the Euler equation shock

$$\eta(t) = \frac{\beta(\alpha\theta(t)K(t-1)^{\alpha} + 1)C(t-1)^{\gamma}}{C(t)^{\gamma}}, \qquad (18)$$

then η satisfies $E_t[\eta(t+1)]=0$, all t. No function of variables dated t-1 or earlier should help predict $\eta(t)$. We can check this condition by estimating a regression of the $\eta(t)$ calculated from the simulation on a constant and variables dated t-1 or earlier. The regression should have an R^2 of zero asymptotically. The model implies that conditional heterosckedasticity is probably present in such a regression when the model has not been backsolved; i.e. $E_t[\eta(t+1)^2|X(t)]\neq$constant, where X(t) is the time t right-hand-side variable vector in the regression. To test correctly for zero R^2 in the equation, therefore, requires that we form a heteroskedasticity-corrected covariance matrix M for the estimated coefficients $\hat{\beta}$ and treat $\hat{\beta}'M^{-1}\hat{\beta}$ as $\chi^2(k)$, where k is the order of the X vector. Without conditional heteroskedasticity, the corresponding test can be formed by treating TR^2 as $\chi^2(k)$, where T is sample size.

The backsolved solution, since it generates $\eta(t)$ directly from a random number generator, satisfies the zero R^2 condition by construction (though checking it is a useful diagnostic for the random number generator). Indeed for the simulation II backsolved solution regressions of $\eta(t)$ on 4 lagged values each of η, C, K, and θ produced $\chi^2(16)=22.8$, which has a marginal significance level of .16. A regression on lagged K and θ alone produced $\chi^2(8)=11.9$, which has a marginal significance level of .16. Results were similar for the backsolved simulation I.

Similar regressions for the loglinear LQ approximate model in simulation I also showed no statistically significant predictability for η. Simulation II for the loglinear LQ method, however, produced $\chi^2(8)=51.0$, which has marginal significance

level well under .01.[7] These regressions do not have large R^2's, but they are large enough that 1000 observations makes them clearly statistically significant. Table 1 below displays the regression of the LQ η on 8 lagged variables. Chart 2 displays a plot of residuals against predicted values from the Table 1 regression. For comparison, Chart 3 displays on the same scale the scatter of residuals against predicted values for the corresponding regression using the backsolved simulation.

TABLE 1

Regression of Euler Equation Disturbance
from LQ Simulation on Lagged K and θ

$$\eta(t) = \underset{(4.4)}{.00478} + \underset{(2.25)}{.00365}\,K(t-1) - \underset{(-.52)}{.00177}\,K(t-2) - \underset{(-1.01)}{.00342}\,K(t-3)$$

$$+ \underset{(1.05)}{.00165}\,K(t-4) - \underset{(-2.80)}{.0148}\,\theta(t-1) - \underset{(-.53)}{.00373}\,\theta(t-2)$$

$$+ \underset{(1.50)}{.00816}\,\theta(t-3) - \underset{(-1.25)}{.00203}\,\theta(t-4)$$

OBSERVATIONS	995	DEGREES OF FREEDOM	986
R**2	.0424	SEE	.00545
DURBIN-WATSON	1.95	$\chi^2(8)$	51.0

Numbers in parenthesis below coefficients are t statistics. χ^2 statistic is heteroskedasticity-corrected as described in text, and tests the null hypothesis that all coefficients other than the constant term are zero.

---

[7]The regression with 16 explanatory variables produced numerical problems in inverting M. However it showed R^2=.0539, implying $\chi^2(16)$=52.7 if we ignore the heteroskedasticity problem. This is also significant at well under the 1% level.

The simulations also generate values for the ν process. Because (15), the true equation generating θ, is loglinear, our loglinear LQ solution method does not distort the distribution of ν at all. The backsolving method, however, is aimed at concentrating all approximation error in the distribution of ν. The histogram of ν in the backsolved solution showed no notable deviation from a normal shape or from the zero mean, standard deviation .10 assumption. A regression of $\nu(t)$ on the same 16 and 8 lagged variables used in the tests for η above produced $\chi^2(16)=27.5$ and $\chi^2(8)=15.8$, which have marginal significance levels of .036 and .045, respectively.[8] There is thus some evidence of predictability of ν in the backsolved simulation, but the effect is much weaker than the predictability of the LQ solution's η.

That the inaccuracy in the LQ simulation displayed here is nontrivial is apparent also in comparing the two regressions in Tables 2 and 3. Both display regressions of C(t) on lagged values of C, K and θ. Regressions of this type might be used to test the random walk hypothesis for consumption. Table 2 shows results for the LQ simulation, Table 3 for the backsolved simulation. While both show firm rejection of the random walk hypothesis (a result which seems to be generic for this model when γ is small), the pattern of estimated coefficients is strikingly different between the two solutions. In one, lagged K is highly significant and lagged theta insignificant, while in the other the reverse is true.

---

[8]These significance levels may be misleading. On the null hypothesis of "accuracy", there is no heteroskedasticity in ν, so the TR^2 statistic applies. These statistics are 24.2 and 14.3, with marginal significance levels of .085 and .075, respectively.

TABLE 2

Consumption Regression, LQ Loglinear Simulation II

$$C(t) = \underset{(-7.8)}{-.0970} + \underset{(19.7)}{1.35}\, C(t-1) - \underset{(-5.6)}{.0203}\, K(t-1) + \underset{(.70)}{.0194}\, \theta(t-1)$$

OBSERVATIONS	999	DEGREES OF FREEDOM	995
R**2	.9976	SEE	.0617
DURBIN-WATSON	2.14		

Numbers in parenthesis below coefficients are t statistics.

TABLE 3

Consumption Regression, Backsolved Simulation II

$$C(t) = \underset{(-3.2)}{-.0458} + \underset{(14.1)}{1.03}\, C(t-1) - \underset{(-.88)}{.0034}\, K(t-1) + \underset{(4.3)}{.0142}\, \theta(t-1)$$

OBSERVATIONS	999	DEGREES OF FREEDOM	995
R**2	.9972	SEE	.0687
DURBIN-WATSON	2.05		

Numbers in parenthesis below coefficients are t statistics.

This example illustrates the fact that high accuracy of simulated solutions for one LQ approximation (here the backsolved simulation) is no guarantee that another (here the direct LQ system simulation) is accurate on the same problem. Not only is there no general rule that "LQ approximate solution is accurate in economic growth models," accuracy of one LQ approximation is no guarantee that another, differing in details of implementation, will also be accurate on the same problem. LQ approximation can be highly accurate, as the simulation I results

showed, but when LQ solutions are presented as characterizing the behavior of a nonlinear model, they should always be supported by accuracy checks. It is easy enough to show the extent to which the original equations and stochastic specifications of the nonlinear model are violated by the approximate simulated solution.

In fact this point applies beyond the context of LQ approximations. All methods of solving continuous parameter, multiple state, nonlinear stochastic equilibrium models are approximate in one way or another. The fact that our solution for the decision rule is approximate, however, is no barrier to our checking the equations and stochastic specifications of the original model against the simulated data from the approximate solution.

IV. Example 2: LQ Savings Problem with Binding Nonnegativity

The problem in this section is behaviorally even simpler than that in section III, but it is inherently nonstationary. Backsolving works quite easily in differentiable problems where the LQ approximate solution provides a good starting point for choice of a decision rule. Here, by contrast, the LQ rule is inherently deficient. Nonetheless backsolved simulations are feasible and possibly useful.

We consider the problem of maximizing

$$E\left[\sum_{t=1}^{\infty}\left(C(t)-.5C(t)^2\right)\beta^t\right] \tag{19}$$

subject to

$$C(t) + K(t) = \alpha K(t-1) + Y(t), \tag{20}$$

$$K(t) \geq 0, \text{ all } t\ . \tag{21}$$

The stochastic process Y is assumed to make $Y(t)>0$, all t with probability one and to make Y i.i.d. across time. C and K are chosen subject to the constraint that C(t) and K(t) can depend only on Y(t) dated t and earlier.

The stochastic Kuhn-Tucker Euler equations for this problem are

$$1 - C(t) = \lambda(t) \tag{22}$$

$$\lambda(t) - \mu(t) = \beta\alpha E_t[\lambda(t+1)] \tag{23}$$

$$\mu(t)K(t)=0 \quad , \tag{24}$$

where λ and μ are the stochastic Lagrange multipliers on (20) and (21), respectively.

Note that C will never exceed one in a solution to this problem. Any candidate solution with $C(t)>1$ at some t can be improved by decreasing C at that date without violating any constraint, because there is no upper bound on K or its rate of growth. If $\beta\alpha\geq 1$, (23) implies that λ_t is a supermartingale (i.e., $E_t\lambda_{t+1}\leq\lambda_t$), since we have structured the problem so $\mu_t\geq 0$, all t. But $\lambda=1-C$ is bounded below by zero, and a supermartingale bounded below converges with probability one. In fact, one can easily see from (23) that when $\alpha\beta>1$ the expectation of $\lambda_{t+s}\to 0$ as $s\to\infty$, so that λ (being strictly positive) converges to 0. That is, C converges to its satiation level of one. Indeed, the same holds for $\alpha\beta=1$ so long as Y(t) does not settle at zero forever, because in that case (23) still guarantees that marginal utility is a supermartingale bounded below, therefore convergent with probability one, and convergence to any value other than zero is impossible. When $\alpha\beta<1$, stationary solutions are possible. Indeed if initial K is not large and the ratio of maximum Y to mean Y is less than $1/\alpha\beta$, then the optimal solution is to consume all the capital in the first period, then set $C(t)\equiv Y(t)$, $K(t)\equiv 0$ thereafter.

Obtaining simulated solutions for this model when $\alpha b \geq 1$ is difficult. Backsolving could, however, form the basis for an iterative approach to the solution. Consider the case of $\alpha b > 1$, and guess that $\mu_t = 0$, all t. Then (23) and the boundary conditions we have discussed imply that C(t) must converge to one, showing stochastic variation along the way, but never equalling or exceeding one unless it hits one exactly and "sticks" there. Furthermore, because when C(t)=1 C(t+s)=1 with probability one for all s>0, C(t)=1 implies $(\alpha-1)K(t)+Y_{min}=1$, where Y_{min} is the lower bound on the support of the distribution of Y(t). (Otherwise, since the probability of an indefinitely long sequence of Y's arbitrarily close to Y_{min} is never zero, the probability of C(t)≡1 forcing K eventually to zero and thereafter becoming infeasible is not zero.) We know that in the LQ problem with no bound on K but with $K(t)\beta^{-.5t} \to 0$ imposed, the optimal policy sets

$$C(t) = (\alpha^2\beta - 1)K(t) + 1 - \frac{(1-\bar{Y})(\alpha^2\beta - 1)}{\alpha - 1} . \qquad (25)$$

When C is far below 1, this is probably a good approximation. On the other hand this formula has C approach 1 when K approaches $(1-\bar{Y})/(\alpha-1)$ rather than when it approaches $(1-Y_{min})/(\alpha-1)$, as we know must be the case in a correct solution.

A reasonable start toward solution, then, might postulate a relation between C and K which is almost exactly (25) for small values of C and K, but which makes C smaller for a given K than is implied by (25) when C and K approach 1 and $(1-Y_{min})/(\alpha-1)$. With such a hypothetical rule in hand, probably parameterized as a flexible nonlinear functional form, one could then generate backsolved simulations. Shocks for equation (23) would be drawn, taking account of the fact that when 1-C(t) is small, the support of the distribution of shocks must be adjusted so that C(t)≤1 with probability one. The resulting simulated time path for C would be

combined with (25) and (20) to find simulated paths for Y and K. The simulated Y paths would then be checked for conformity to the originally postulated distribution and i.i.d. assumption.[9]

Note that in this nonstationary problem the invertibility and stability problems do not arise directly. By choosing the modification of (25) so that it guarantees $K(t) \geq (1-Y_{min})/(\alpha-1)$ when $C(t)=1$, we enforce directly a terminal condition which is the analogue of the stability condition in a model with stationary solutions. Imposing this terminal condition may require strange behavior of Y as C nears 1, and this is the analogue of the unstable Euler equation solutions which have to be ruled out in stationary models.

V. Iteration Toward Exact Solution

The methods we are discussing here allow computation of simulated solutions to an optimization problem or an equilibrium model given arbitrary "decision rules". We have suggested using rules which match a solution to a linearized version of the models stochastic equations in the neighborhood of a steady state. But this is only a reasonable starting point. More generally, the decision rule (or in an equilibrium model decision rules and pricing functions) can be parametrized flexibly and the space of solutions generated

---

[9] I have experimented with this approach. Two or three simple guesses for a modification of (25) failed to produce a solution with good behavior for Y when C neared (1), however. To make the procedure work will apparently require systematic iteration over the functional form of the substitute for (25), minimizing a measure of the discrepancy between desired and actual properties of the Y simulations.

as the rules' parameters are varied can be explored. A numerical optimization program can be used to minimize the distance between an hypothesized distribution for the exogenous stochastic shocks in the model and the distribution generated by the backsolved simulations.

Of course the same idea could be applied to forward simulations. The difference is that forward simulations would have exogenous shocks satisfying the assumed distribution by construction, but failing to satisfy the Euler equations. Thus for forward solutions the appropriate fit criterion would be minus some measure of the predictability of the Euler equation shocks. (Marcet [1988] has applied a similar idea successfully, though he flexibly parameterizes conditional expectation functions rather than decision rules or pricing functions.)

VI. Estimation

These methods make it possible to simulate models of size comparable to common macroeconometric models. I have used it to solve and simulate several versions of a model of money, nominal interest rate, price level, capital and output, with demand for money motivated by transactions costs. The model treats investment and consumption as distinct goods. While the reported model [1989] is still first-order, it is capable of roughly matching many aspects of the MAR for money, interest rates, output and prices in the observed data.

This raises the prospect of using these simulation techniques to estimate non-LQ dynamic models. The principle is simple -- one adjusts parameters of the model until some set of functions of the simulated sample paths matches as closely as possible the corresponding set of functions of the observed historical data. Ingram and Lee [1989] discuss the classical asymptotic econometric

theory of such estimators and Kwan [1989] provides a way to give such results a Bayesian interpretation.

Note that in the special case where the theoretical model is at the same discrete time unit as the data and contains just as many unobservable stochastic exogenous variables as there are observable variables, estimation by simulation may not be necessary. In this special case, the same set of equations which allows simulation of artificial data given computer-generated artificial random disturbances can be solved to find the random disturbances implied by the observed historical data. With an hypothesis on the distribution of the disturbances, it then becomes possible to formulate a likelihood function, allowing standard maximum likelihood estimation techniques.

However it will be quite common for the number of underlying behavioral shocks to exceed the number of observed variables, or for the theoretical model naturally to be treated as at a finer time scale than the observed data. In these cases estimation by simulation can substitute for numerical analytic integration which would otherwise be needed to form a marginal distribution for the observed data.

REFERENCES

Ingram, Beth and Bong-Soo Lee [1989]. "Estimation by Simulation", forthcoming, Journal of Econometrics.

Prescott, Edward and Finn Kydland [1982]. "Time to Build and Aggregate Fluctuations," Econometrica, v.50, no.6, 1345-1370.

Kushner, Harold [1971]. Introduction to Stochastic Control, Holt, Rinehart and Winston.

Kwan, Yeum-Keung [1988]. "The Asymptotic Properties of Quasi Posterior Density and a Limited Information Bayesian Procedure", processed, U. of Minnesota.

Marcet, Albert [1988]. "Solution Method", paper presented at September 1988 conference in Minneapolis and available from Marcet at Carnegie-Mellon University.

Novales, Alfonso [1984]. "A Stochastic Monetary Equilibrium Model of the Interest Rate," to appear, Econometrica.

Schaefer, Helmut [1966]. Topological Vector Spaces, Macmillan.

Sims, Christopher [1989]. "Models and their Uses", forthcoming, American Journal of Agricultural Economics.

CHART 1.I

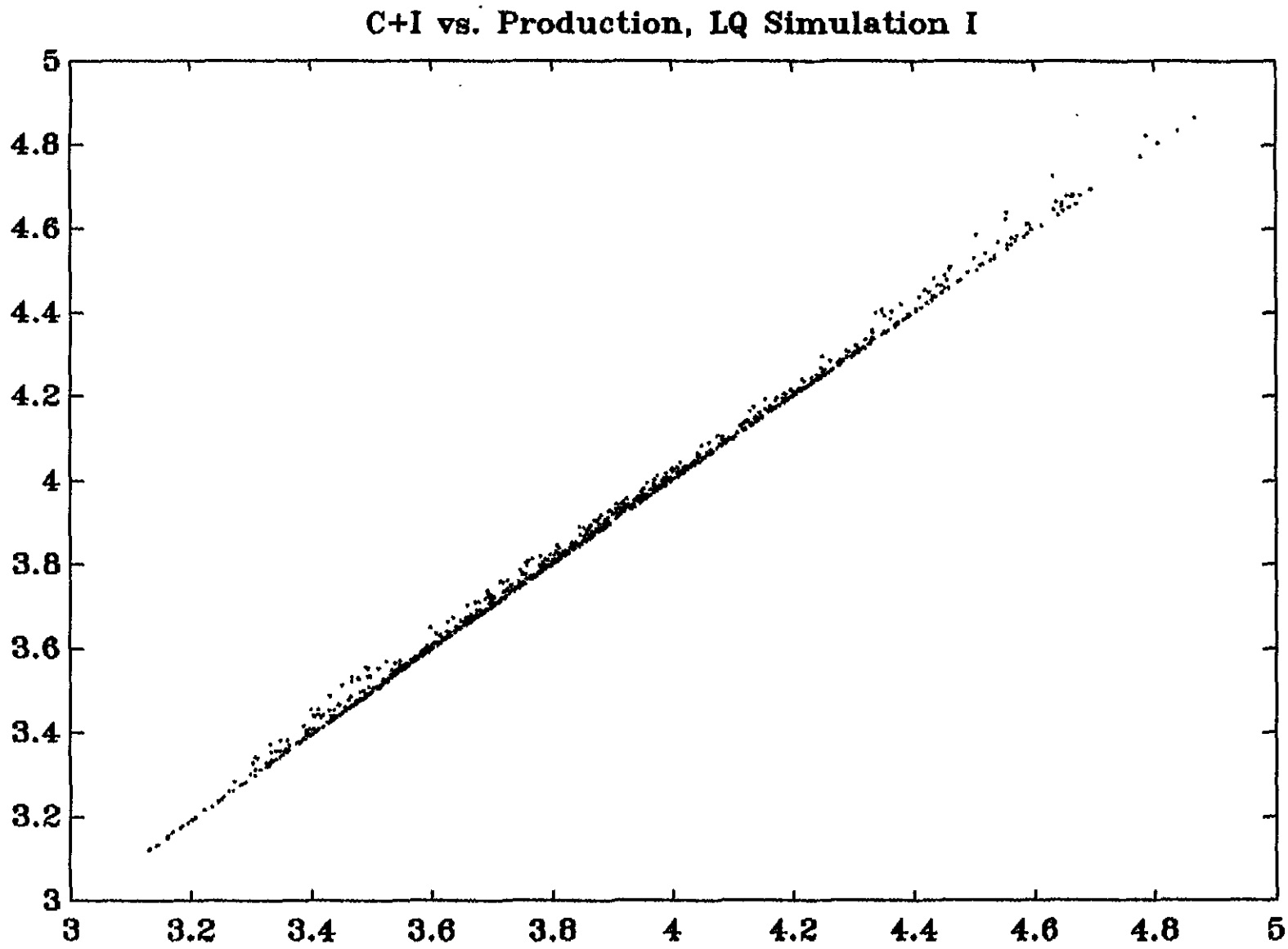

C+I vs. Production, LQ Simulation I
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
3
3.2
3.4
3.6
3.8
4
4.2
4.4
4.6
4.8
5

CHART 1.II

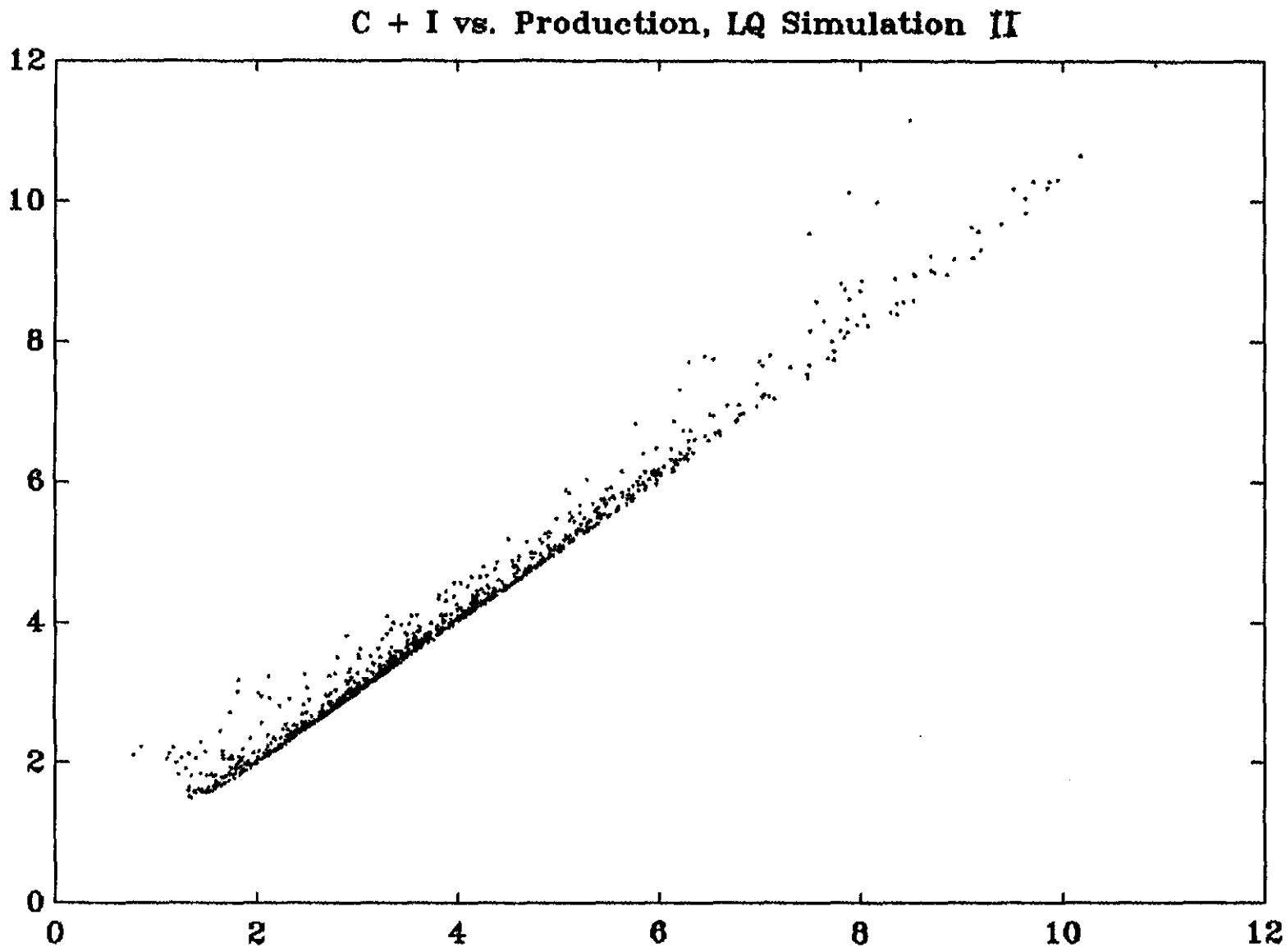


CHART 2

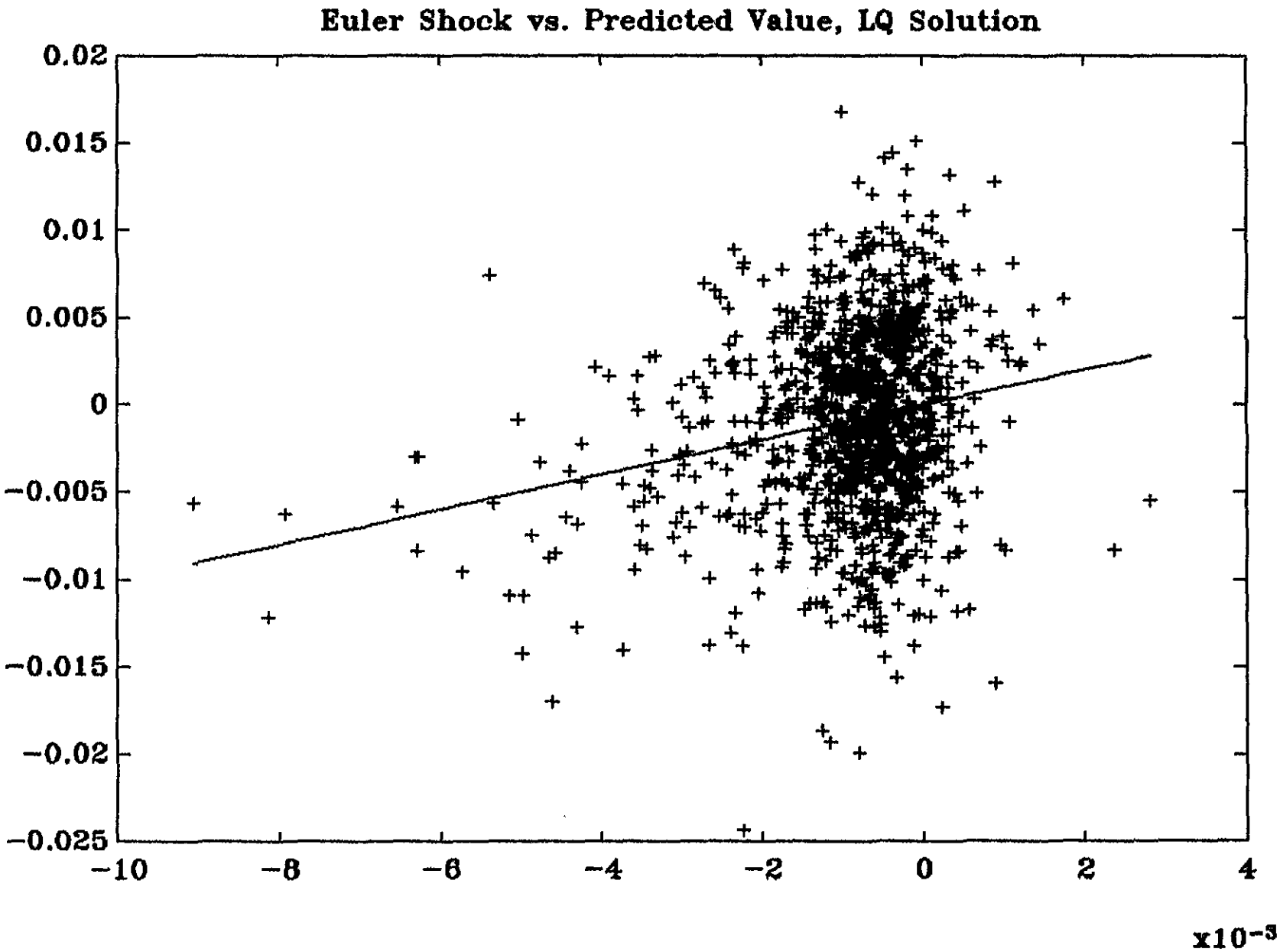
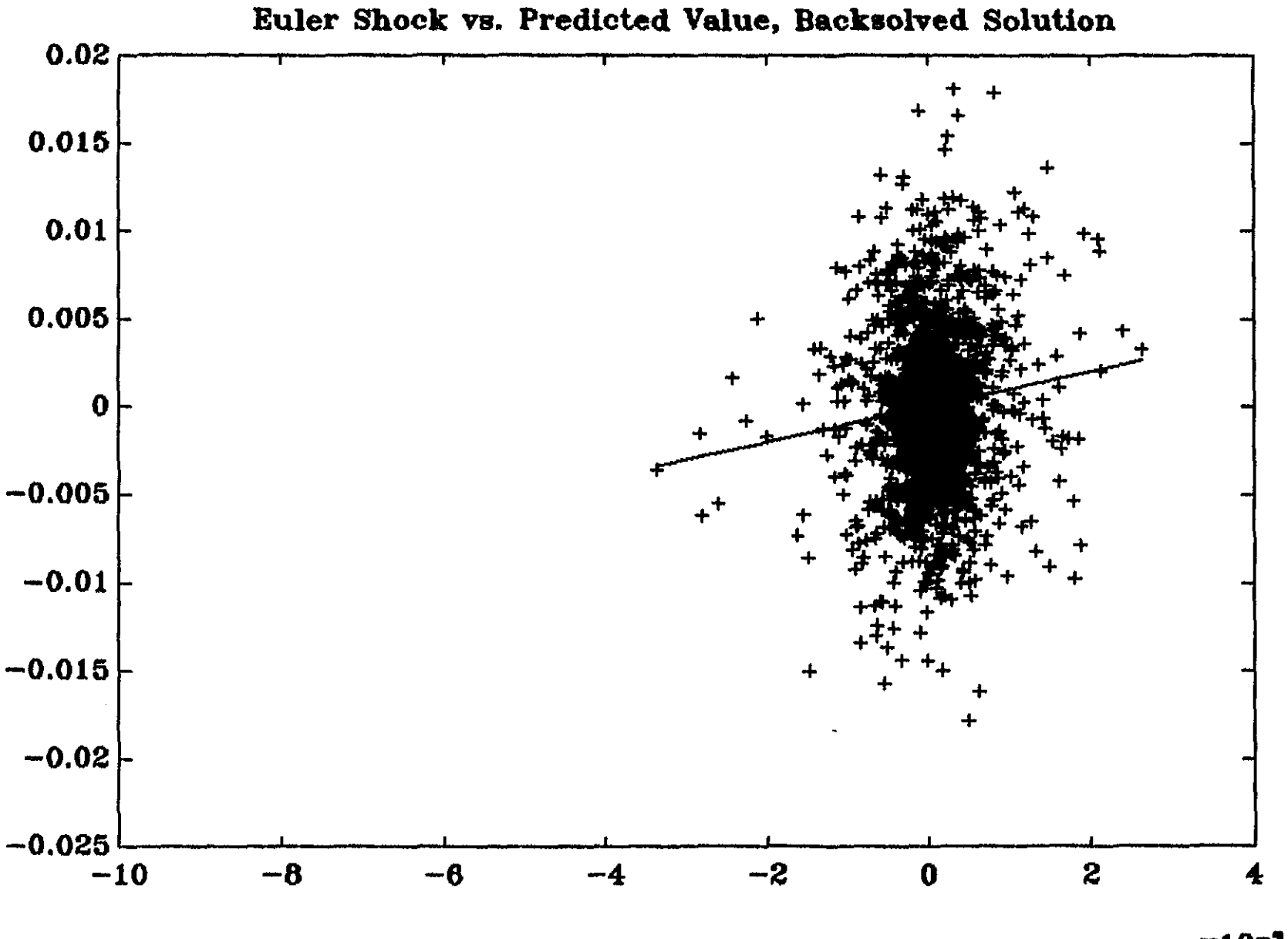

Euler Shock vs. Predicted Value, LQ Solution
0.02
0.015
0.01
0.005
0
−0.005
−0.01
−0.015
−0.02
−0.025
−10
−8
−6
−4
−2
0
2
4
x10−3

CHART 3


Euler Shock vs. Predicted Value, Backsolved Solution
0.02
0.015
0.01
0.005
0
−0.005
−0.01
−0.015
−0.02
−0.025
−10
−8
−6
−4
−2
0
2
4
x10−3