

Perturbation Methods for General Dynamic Stochastic Models

HE-HUI JIN*

KENNETH L. JUDD

DEPARTMENT OF ECONOMICS

HOOVER INSTITUTION

STANFORD UNIVERSITY

STANFORD, CA 94305

STANFORD, CA 94305

JUDD@HOOVER.STANFORD.EDU

December, 2002

ABSTRACT. We describe a general Taylor series method for computing asymptotically valid approximations to deterministic and stochastic rational expectations models near the deterministic steady state. Contrary to conventional wisdom, the higher-order terms are conceptually no more difficult to compute than the conventional deterministic linear approximations. We display the solvability conditions for second- and third-order approximations and show how to compute the solvability conditions in general. We use an implicit function theorem to prove a local existence theorem for the general stochastic model given existence of the degenerate deterministic model. We describe an algorithm which takes as input the equilibrium equations and an integer k , and computes the order k Taylor series expansion along with diagnostic indices indicating the quality of the approximation. We apply this algorithm to some multidimensional problems and show that the resulting nonlinear approximations are far superior to linear approximations.

*We thank Mordecai Kurz and Chris Sims for valuable comments.

Economists are using increasingly complex dynamic stochastic models and need more powerful and reliable computational methods for their analysis. We describe a general perturbation method for computing asymptotically valid approximations to general stochastic rational expectations models based on their deterministic steady states. These approximations go beyond the normal “linearize around the steady state” approximations by adding both higher-order terms and deviations from certainty equivalence. The higher-order terms and corrections for risk will likely improve the accuracy of the approximations and their useful range. Also, some questions, such as welfare effects of security markets, can be answered only with higher-order approximations; see Judd and Guu (2001) for models where higher-order terms are essential. Contrary to conventional wisdom, these higher-order terms are no more difficult to compute than the conventional deterministic linear approximations; in fact, they are conceptually easier. However, we show that one cannot just assume that the higher-order terms create a better approximation. We examine the relevant implicit function theorems that justify perturbation methods in some cases and point out cases where perturbation methods may fail. Since perturbation methods are not perfectly reliable, we also present diagnostic procedures which will indicate the reliability of any specific approximation. Since the diagnostic procedures consume little computational effort compared with the construction of the approximation, they produce critical information at little cost.

Linearizations methods for dynamic models have been a workhorse of macroeconomic analysis. Magill (1977) showed how to compute a linear approximation around deterministic steady states and apply them to approximate spectral properties of stochastic models. Kydland and Prescott (1982) applied a special case of the Magill method to a real business cycle model. However, the approximations in Magill, and Kydland and Prescott were just linear approximations of the deterministic model applied to stochastic models; they ignored higher-order terms and were certainty

equivalent approximations, that is, variance had no impact on decision rules. The motivating intuition was also specific to the case of linear, certainty equivalent, approximations. Kydland and Prescott (1982) motivated their procedure by replacing the nonlinear law of motion with a linear law of motion and replacing the nonlinear payoff function with a quadratic approximation, and then applying linear-quadratic dynamic programming methods to the approximate model. This motivation gives the impression that it is not easy to compute higher-order approximations, particularly since computing the first-order terms requires solving a quadratic matrix equation. In fact, Marcet(1994) dismissed the possibility that higher-order approximations be computed, stating that “perturbation methods of order higher than one are considerably more complicated than the traditional linear-quadratic case ...”

Furthermore, little effort has been made to determine the conditions under which certainty equivalent linearizations are valid. Linearization methods are typically used in an application without examining whether they are valid in that case. This raises questions about many of the applications, particularly since the conventional linearization approach sometimes produces clearly erroneous results. For example, Tesar (1995) uses the standard Kydland-Prescott method and found an example where completing asset markets will make all agents worse off. This result violates general equilibrium theory and can only be attributed to the numerical method used. Kim and Kim (forthcoming) show that this will often occur in simple stochastic models. Below we will present a portfolio-like example which shows that casual applications of higher-order procedures (such as those advocated by Sims, 2002, and Campbell and Viciara, 2002) can easily produce nonsensical answers. These examples emphasize two important points. First, more flexible, robust, and accurate methods based on sound mathematical principles are needed. Second, we cannot blindly accept the results of a Taylor series approximation but need ways to test an approximation’s reliability. This paper addresses both issues.

We will show that it is practical to compute higher-order terms to the multivariate Taylor series approximation based at the deterministic steady state. The basic fact shown below is that all the higher-order terms of the Taylor series expansion, even in the stochastic multidimensional case, are solutions to linear problems once one computes the first-order terms. This implies that the higher-order terms are *easier* to compute in the sense that linear problems are conceptually less complex. In previous papers, Judd and Guu (1993, 1997) examined perturbation methods for deterministic, continuous- and discrete-time growth models in one capital stock, and stochastic growth models in continuous time with one state. They find that the high-order approximations can be used to compute highly accurate approximations which avoid the certainty equivalence property of the standard linearization method. Judd and Gaspar (1997) described perturbation methods for multidimensional stochastic models in continuous time, and produced Fortran computer code for fourth-order expansions. Judd (1998) presented the general method for deterministic discrete-time models and presented a discrete-time stochastic example indicating the critical adjustments necessary to move from continuous time to discrete time. In particular, the natural perturbation parameter is the instantaneous variance in the continuous-time case, but the standard deviation is the natural perturbation parameter for discrete-time stochastic models. The reader is referred to these papers and their mathematical sources for key definitions and introductions to these methods. In this paper, we will outline how these methods can be adapted to handle more general rational expectations problems.

There has recently been an increase in the interest in higher-order approximation methods. Collard and Juillard (2001a) computed a higher-order perturbation approximation of an asset-pricing model and Collard and Juillard (2001b). Chen and Zadrozny (forthcoming) computed higher-order approximations for a family of optimal control problems. Kim and Kim (forthcoming) applied second-order ap-

proximation methods to welfare questions in international trade. Sims (2000) and Grohe-Schmidt and Uribe (2002) have generalized Judd (1998), Judd and Gaspar (1997), and Judd and Guu (1993) by examining second-order approximations of multidimensional discrete-time models.

The first key step is to express the problem formally as two different kinds of perturbation problems and apply the appropriate implicit function theorems. Even though we are applying ideas from implicit function theory, there are unique difficulties which arise in stochastic dynamic models. Perturbation methods revolve around *solvability conditions*, that is, conditions which guarantee a unique solution to terms in an asymptotic expansion. We display the solvability conditions for Taylor series expansions of arbitrary orders for both deterministic and stochastic problems, showing that they reduce to the invertibility of a series of matrices. The implicit function theorem for the deterministic problem is straightforward, but the stochastic components produce novel problems. We give an example where a casual approach will produce a nonsensical result. We use an implicit function theorem to prove a local existence theorem for the general stochastic model given existence of the degenerate deterministic model. This is a nontrivial step and an important one since it is easy for economists to specify models which lack a local existence theorem justifying perturbation methods.

We then describe an algorithm which takes as input the equilibrium equations and an integer k , and computes the order k Taylor series expansion along with diagnostic indices indicating the quality of the approximation. We apply this algorithm to some multidimensional problems and show that the resulting nonlinear approximations are far superior to linear approximations over a large range of states. We also emphasize the importance of error estimation along with computation of the approximation.

1. A PERTURBATION APPROACH TO THE GENERAL RATIONAL EXPECTATIONS PROBLEM

We examine general stochastic problems of the form

$$\begin{aligned} 0 &= E \{ g^\mu(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z) | x_t \}, \mu = 1, \dots, m \\ x_{t+1}^i &= F^i(x_t, y_t, \varepsilon z_t), i = 1, \dots, n \end{aligned} \quad (1)$$

where $x_t \in \mathbb{R}^n$ are the predetermined variables at the beginning of time t , such as capital stock, lagged variables and productivity, $y_t \in \mathbb{R}^m$ are the endogenous variables at time t , such as consumption, labor supply and prices, and $F^i(x_t, y_t, \varepsilon z_t) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^s \rightarrow \mathbb{R}$, $i = 1, \dots, n$ is the law of motion for x^i , and

$$g^\mu(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}, \mu = 1, \dots, q$$

are the equations defining equilibrium, including Euler equations and market clearing conditions. The scalar ε is a scaling parameter for the disturbance terms z . We assume that the components of z are i.i.d. with mean zero and unit variance, making ε the common standard deviation. Since correlation and heteroscedasticity can be built into the function g , we can do this without loss of generality. Different values for ε represent economies with different levels of uncertainty. The objective is to find some equilibrium rule, $Y(x, \varepsilon)$, such that in the ε -economy the endogenous and predetermined variables satisfy

$$y_t = Y(x_t, \varepsilon)$$

This implies that $Y(x, \varepsilon)$ must satisfy the functional equation

$$E \{ g^\mu(x, Y(x, \varepsilon), F(x, Y(x, \varepsilon), \varepsilon z), Y(F(x, Y(x, \varepsilon), \varepsilon z), \varepsilon), \varepsilon z) | x \} = 0 \quad (2)$$

Our perturbation method will approximate $Y(x, \varepsilon)$ with a polynomial in (x, ε) in a neighborhood of the deterministic steady state. The deterministic steady state is

the solution to

$$\begin{aligned} 0 &= g(x_*, y_*, x_*, y_*, 0) \\ x_* &= F(x_*, y_*, 0) \end{aligned} \tag{3}$$

The steady state y_* is the first step in approximating $Y(x, \varepsilon)$ since $Y(x_*, 0) = y_*$. The task is to find the derivatives of $Y(x, \varepsilon)$ with respect to x and ε at the deterministic steady state, and use that information to construct a degree k Taylor series approximation of $Y(x, \varepsilon)$, such as in

$$\begin{aligned} Y(x_* + v, \varepsilon) &\doteq y_* + Y_x(x_*, 0)v + \varepsilon Y_\varepsilon(x_*, 0) \\ &+ v^\top Y_{xx}(x_*, 0)v + \varepsilon Y_{x\varepsilon}(x_*, 0)v + \varepsilon^2 Y_{\varepsilon\varepsilon}(x_*, 0) \\ &+ \dots \\ &+ o\left(\varepsilon^k + \|v\|^k\right) \end{aligned} \tag{4}$$

If Y is analytic in the neighborhood of $(x_*, 0)$ then this series has an infinite number of terms and it is locally convergent. The objective is also to be able to use the Taylor series approximation in simulations of the nonlinear model and be able to produce uniformly valid approximations of the long-run and short-run behavior of the true nonlinear model. This is a long list of requirements but we will develop diagnostics to check out the performance of our Taylor series approximations.

Equation (1) includes a broad range of dynamic stochastic models, but does leave out some models. For example, models with intertemporally nonseparable preferences, like those in Epstein-Zin (1989), are functional equations and do not obviously reduce to a dynamic system in \mathbb{R}^m . However, with modest modifications, our methods can be applied to any problem of the form in (2), a larger set of problems than those expressible as (1). We also assume that any solution to (3) is locally unique. This rules out many interesting models, particularly models with portfolio choices and

models where income distribution may matter. Portfolio problems can probably be handled with dynamic extensions of Judd and Guu (2001), and income distribution problems can probably be handled by application of the center manifold theorem, but we leave these developments for later work.

Computing and evaluating the approximation in (4) is accomplished in five steps. The first is to solve (3) for steady state values (x_*, y_*) . This is presumably accomplished by applying some nonlinear equation solver to (3) and will not be further discussed here. The second is to compute the linear approximation terms, $Y_x(x_*, 0)$. This is done by analyzing the deterministic system formed by setting $\varepsilon = 0$ in (1) to create the perfect foresight system

$$\begin{aligned} 0 &= g^\mu(x_t, y_t, x_{t+1}, y_{t+1}, 0) \\ x_{t+1}^i &= F^i(x_t, y_t, 0) \end{aligned} \tag{6}$$

Computing the linear terms is a standard computation, solvable by a variety of techniques. See the literature on linear rational expectations models for solution methods (Anderson, et al. , 1996, is a survey of this literature); we will not discuss this step further.

This paper is concerned with the next three steps. Third, we compute the higher-order deterministic terms; that is, we compute perturbations of $Y(x, 0)$ in the x directions. Formally, we want to compute $\frac{\partial}{\partial x^k} Y(x_*, 0)$, $k = 1, 2, \dots$. This produces the Taylor series approximation for the deterministic problem

$$Y(x, 0) \doteq y_* + Y_x(x_*, 0)(x - x_*) + (x - x_*)^\top Y_{xx}(x_*, 0)(x - x_*) + \dots \tag{7}$$

for the solution to (6).

Fourth, with the Taylor series for $Y(x, 0)$ in hand, we examine the general stochastic problem $Y(x, \varepsilon)$. We use the expansion (7) of the deterministic problem to compute the ε derivatives, $\left(\frac{\partial}{\partial \varepsilon}\right)^\ell \left(\frac{\partial}{\partial x}\right)^k Y(x_*, 0)$. More generally, we show that how

to take a solution of $Y(x, 0)$ and use it to construct a solution to $Y(x, \varepsilon)$ for small ε . This last step raises the possibility that we have an approximation which is not just locally valid around the deterministic steady state point $(x_*, 0)$ but instead around a large portion of the stable manifold defined by $Y(x, 0)$.

This four-stage approach is the proper procedure since each step requires solutions from the previous steps. Also, by separating the stochastic step from the deterministic steps we see the main point that we can perturb around the deterministic stable manifold, not just the deterministic steady state.

Before we accept the resulting candidate Taylor series, we must test its reliability. Let $\hat{Y}(x, \varepsilon)$ be the computed finite order Taylor series we have computed. We evaluate it by computing

$$E \left\{ \tilde{g}^\mu(x, \hat{Y}(x, \varepsilon), F(x, \hat{Y}(x, \varepsilon), \varepsilon z), \hat{Y}(F(x, \hat{Y}(x, \varepsilon), \varepsilon z), \varepsilon), \varepsilon z) | x \right\} = 0$$

for a range of values of (x, ε) that we want to use, where $\tilde{g}^\mu(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z)$ is a unit-free version of $g^\mu(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z)$. That is, each component of $E \{ \tilde{g}^\mu \}$ is transformed so that any deviation from zero represent a relative error. For example, if one component of g^μ is supply equals demand then the corresponding component of \tilde{g}^μ will express excess demand as a fraction of total demand and any deviation of $E \{ \tilde{g}^\mu(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z) | x_t \}$ from zero represents the relative error in the supply equals demand condition. If these relative errors are sufficiently small then we will accept $\hat{Y}(x, \varepsilon)$. This last step is critical since Taylor series expansions have only a finite range of validity and we have no a priori way of knowing the range of validity.

Before continuing, we warn the reader of the nontrivial notational challenge which awaits him in the sections below where we develop the theoretical properties of our perturbation method and present the formal justification of our algorithm. After being introduced to tensor notation and its application to multivariate stochastic control, the reader may decide that this approach is far too burdensome to be of

value. If one had to go through these manipulations for each and every application, we might agree. Fortunately, all of the algebra discussed below has been automated, executing *all* the necessary computations, including analytic derivatives and error indices, and produce the Taylor series approximation discussed below. This will relieve the user of executing all the algebra we discuss below.

2. MULTIDIMENSIONAL COMPARATIVE STATICS AND TENSOR NOTATION

We first review the tensor notation necessary to efficiently express the critical multivariate formulas. We will follow the tensor notation conventions used in mathematics (see, for example, Bishop and Goldberg, 1981, and Misner et al., 1973) and statistics (see McCullagh, 1987), and use standard adaptations to deal with idiosyncratic features of rational expectations models. We then review the implicit function theorem, and higher-order applications of the implicit function theorem.

2.1. Tensor Notation. Multidimensional perturbation problems use the multidimensional chain rule. Unfortunately, the chain rule in R^n produces a complex sequence of summations, and conventional notation becomes unwieldy. The *Einstein summation notation* for tensors and its adaptations will give us a natural way to address the notational problems.¹ Tensor notation is a powerful way of dealing with multidimensional collections of numbers and operations involving them. We will present the elements of tensor notation necessary for our task; see Judd (1998) for more discussion.

Suppose that a_i is a collection of numbers indexed by $i = 1, \dots, n$, and that x^i is a singly indexed collection of real numbers. Then

$$a_i x^i \equiv \sum_i a_i x^i.$$

¹The dubious reader should try to read the formulas in Bensoussan(1988) where conventional notation is used.

is the tensor notation for the inner product of the vectors represented by a_i and x^i . This notation is desirable since it eliminates the unnecessary Σ symbol. Similarly suppose that a_{ij} is a collection of numbers indexed by $i, j = 1, \dots, n$. Then

$$a_{ij} x^i y^j \equiv \sum_i \sum_j a_{ij} x^i y^j.$$

is the tensor notation for a quadratic form. Similarly, $a_j^i x_i y^j$ is the quadratic form of the matrix a_j^i with the vectors x and y , and the expression $z_j = a_j^i x_i$ can also be thought of as a matrix multiplying a vector. We will often make use of the Kronecker tensor, which is defined as

$$\delta_j^i \equiv \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

and is a representation of the identity matrix. δ_β^α , δ_J^I , etc., are similarly defined.

More formally, we let x^i denote any vector in R^n and let a_i denote any element in the dual space of R^n , that is, a linear map on vectors x^i in \mathbb{R}^n . Of course, the dual space of \mathbb{R}^n is \mathbb{R}^n . However, it is useful in tensor algebra to keep the distinction between a vector and an element in the dual space.

In general, $a_{j_1, j_2, \dots, j_m}^{i_1, i_2, \dots, i_\ell}$ is a $\ell - m$ tensor, a set of numbers indexed by ℓ superscripts and m subscripts. It can be thought of as a scalar-valued multilinear map on $(\mathbb{R}^n)^\ell \times (\mathbb{R}^n)^m$. This generalizes the idea that matrices are bilinear maps on $(\mathbb{R}^n)^2$. The summation convention becomes particularly useful for higher-order tensors. For example, in \mathbb{R}^n ,

$$c_{j_3, i_4}^{i_3, j_4} = a_{j_1, j_2, j_3}^{i_1, i_2, i_3} b_{i_1, i_2, i_4}^{j_1, j_2, j_4} \equiv \sum_{i_1=1}^n \sum_{i_2=1}^n \sum_{j_1=1}^n \sum_{j_2=1}^n a_{j_1, j_2, j_3}^{i_1, i_2, i_3} b_{i_1, i_2, i_4}^{j_1, j_2, j_4}$$

In our applications if $f : R \rightarrow \mathbb{R}^m$, then f_j will be the derivative of f with respect to x^j . The following equation expresses the multivariate Taylor expansion in tensor notation:

$$f(x^0 + v) = f(x^0) + f_i(x^0)v^i + \frac{1}{2}f_{ij}(x^0)v^i v^j + \frac{1}{3!}f_{ijk}(x^0)v^i v^j v^k + \dots,$$

where $f_i \equiv \frac{\partial f}{\partial x^i}(x^0)$, $f_{ij} \equiv \frac{\partial^2 f}{\partial x^i \partial x^j}(x^0)$, etc. More generally, if $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, then f_j^i will be the derivative of the i 'th component of f with respect to x^j . We will make extensive use of the multivariate chain rule. If $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^m \rightarrow \mathbb{R}^\ell$ and $h(x) = g(f(x))$, then $h : \mathbb{R}^n \rightarrow \mathbb{R}^\ell$, and the Jacobian of h is

$$h_j^i \equiv \frac{\partial h^i}{\partial x^j} = g_\ell^i f_j^\ell.$$

Furthermore the 1-2 tensor of second-order derivatives is

$$h_{jk}^i \equiv \frac{\partial^2 h^i}{\partial x^j \partial x^k} = g_{\ell m}^i f_k^m f_j^\ell + g_\ell^i f_{jk}^\ell.$$

This can be continued to express arbitrary derivatives.

Equation (1) is based on

$$g(x, y, z, w, \varepsilon) : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \times \mathbb{R}^{n_4} \times \mathbb{R} \rightarrow \mathbb{R}$$

which is a function of five groups of variables. Our perturbation analysis needs to distinguish among derivatives with respect to different subsets of the variables. We will use the standard device of letting different indices denote the differentiation with respect to different sets of variables. For example, in general relativity theory, one typically uses Latin letters, a, b, c, \dots , to represent summation over the three space coordinates (x, y, z) and Greek letters, μ, ν, ρ, \dots , to represent summation over space-time coordinates (x, y, z, t) . We apply this practice here to distinguish among x, y, z , and w . Specifically, the derivatives of g with respect to x^i will be denoted by lower case Latin letters as in

$$g_i^\mu(x, y, z, w, \varepsilon) = \frac{\partial}{\partial x^i} (g^\mu(x, y, z, w, \varepsilon))$$

This implies that

$$g_j^\mu(x, y, z, w, \varepsilon) = \frac{\partial}{\partial x^j} (g^\mu(x, y, z, w, \varepsilon))$$

denotes the partial derivative of $g^\mu(x, y, z, w)$ w.r.t. the j 'th component of x . In general, g with a subscript from $\{i, j, k, \dots\}$ denotes the vector of derivatives of $g^\mu(x, y, z, w, \varepsilon)$ with respect to the components of x . We use different indices to denote derivatives with respect to components of y . In particular, we will use lower case Greek letters to index components of y and define

$$\begin{aligned} g_\alpha(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial y^\alpha} (g(x, y, z, w, \varepsilon)) \\ g_\beta(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial y^\beta} (g(x, y, z, w, \varepsilon)) \end{aligned}$$

Derivatives of $g^\mu(x, y, z, w)$ w.r.t. components of z will use capitalized Latin letters, as in

$$\begin{aligned} g_I^\mu(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial z^I} (g^\mu(x, y, z, w, \varepsilon)) \\ g_J^\mu(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial z^J} (g^\mu(x, y, z, w, \varepsilon)) \end{aligned}$$

and derivatives of $g^\mu(x, y, z, w)$ w.r.t. components of w will use capitalized Greek letters, as in

$$\begin{aligned} g_A^\mu(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial w^A} (g^\mu(x, y, z, w, \varepsilon)) \\ g_\Gamma^\mu(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial w^\Gamma} (g^\mu(x, y, z, w, \varepsilon)) \\ g_\Delta^\mu(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial w^\Delta} (g^\mu(x, y, z, w, \varepsilon)) \end{aligned}$$

The distinction holds only for subscripts. For example, the notation y^i will denote the same vector as would y^I .

Since ε is a scalar, the derivatives w.r.t. ε will be denoted in the standard manner,

$$\begin{aligned} g_\varepsilon^\mu(x, y, z, w, \varepsilon) &= \frac{\partial}{\partial \varepsilon} (g^\mu(x, y, z, w, \varepsilon)) \\ g_{\varepsilon\varepsilon}^\mu(x, y, z, w, \varepsilon) &= \frac{\partial^2}{\partial \varepsilon^2} (g^\mu(x, y, z, w, \varepsilon)) \end{aligned}$$

Combinations of indices represent cross derivatives. For example,

$$g_{\alpha i}^{\mu}(x, y, z, w, \varepsilon) = \frac{\partial}{\partial x^i} \frac{\partial}{\partial y^{\alpha}} (g^{\mu}(x, y, z, w, \varepsilon))$$

We will often want the composite gradient of $g^{\mu}(x, y, z, w, \varepsilon)$ consisting of the derivatives with respect to (y, z, w) . We let $\aleph = \{\alpha, I, A\}$ denote the set of all indices for (y, z, w) and use the notation $g_{\aleph}^{\mu}(x, y, z, w, \varepsilon)$ to represent the derivatives as in

$$g_{\aleph}^{\mu}(x, y, z, w, \varepsilon) = \left(g_{\alpha}^{\mu}(x, y, z, w, \varepsilon), g_I^{\mu}(x, y, z, w, \varepsilon), g_A^{\mu}(x, y, z, w, \varepsilon) \right)$$

2.2. Implicit Function Theorem and Solvability Conditions. The general implicit function theorem for \mathbb{R}^n is the critical tool motivating our computations, even though it is not directly applicable to our problems. A review of the Implicit Function Theorem highlights the critical issues we will face. Let $H(x, y) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, and suppose $H(x_0, y_0) = 0$. Let

$$H^{\mu}(x, h(x)) = 0 \tag{8}$$

implicitly define $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ for x near x_0 . In particular, we know $h(x_0) = y_0$. Taking the derivative of (8) with respect to x^i shows

$$H_i^{\mu}(x, h(x)) + H_{\alpha}^{\mu}(x, h(x)) h_i^{\alpha}(x) = 0$$

which, at $x = x_0$, implies

$$\frac{\partial Y^{\alpha}}{\partial x^i}(x_0) = h_i^{\alpha}(x_0) = -\tilde{H}_{\mu}^{\alpha} H_i^{\mu}$$

where \tilde{H}_{μ}^{α} is the tensor (matrix) satisfying

$$\tilde{H}_{\mu}^{\alpha}(H_{\beta}^{\mu}(x_0, y_0)) = \delta_{\beta}^{\alpha}$$

The tensor h_i^{α} is the comparative static matrix which expresses how components of $h(x)$ change as we move x away from x_0 . The solution h_i^{α} exists as long as \tilde{H}_q^{α} , the

inverse matrix of $H_\alpha^q(x_0, y_0)$, exists. Therefore, the solvability condition for h_i^α is the nonsingularity of $H_\alpha^q(x_0, y_0)$. If $H_\alpha^q(x_0, y_0)$ is invertible, then the linear approximation of $h(x)$ based at $x = x_0$ is

$$h^\alpha(x_0 + v) \doteq h^\alpha(x_0) + h_i^\alpha v^i.$$

which is just the beginning of the multivariate Taylor series approximation of $h(x)$ based at x_0 .

We often need to go beyond the first-order expression of and construct higher-order terms of the Taylor series expansion. To do this in a clean and compact fashion, we need tensor notation. Taking another derivative of (8) w.r.t. x^j implies

$$H_{ij}(x, h(x)) + H_{\alpha j}^\mu(x, h(x))h_i^\alpha(x)h_j^\beta(x) + H_\alpha^\mu(x, h(x))h_{ij}^\alpha(x) = 0$$

which, at $x = x_0$, implies

$$h_{ij}^\alpha(x_0) = -\tilde{H}_q^\alpha \left(H_{ij} + H_{\beta j}^q h_i^\beta h_j^\alpha \right) \quad (9)$$

Again, we see that the second-order coefficients are given by (9) as long as $H_\alpha^\mu(x_0, y_0)$ is invertible. Further differentiation shows that at each stage the critical solvability condition is the same, the invertibility of $H_\alpha^\mu(x_0, y_0)$. Therefore, we can continue to solve for terms in a Taylor series expansion as long as H has the necessary derivatives. We will compute the solvability conditions for the dynamic perturbation problem, and find that they differ from this in that the k 'th order solvability condition depends on k .

3. A SIMPLE EXAMPLE

We first illustrate regular perturbation in the context of the basic rational expectations equations in a simple optimal growth model. This will help us see through the complexity of the multidimensional case and also show why we use ε as a perturbation variable as opposed to the variance ε^2 which is the perturbation variable in

continuous-time asymptotic methods, such as in Fleming, Judd and Guu (1993), and Gaspar and Judd (1997).

Consider the simple stochastic optimal growth problems indexed by ε

$$\begin{aligned} \max_{c_t} \quad & \sum_{t=0}^{\infty} \beta^t u(c_t) \\ \text{s.t.} \quad & k_{t+1} = F(k_t - c_t)(1 + \varepsilon z_t) \end{aligned} \quad (10)$$

where the z_t are i.i.d. with unit variance, and ε is a parameter expressing the standard deviation. The solution of the deterministic case, $\varepsilon = 0$, can be expressed as a policy function, $C(k)$, satisfying the Euler equation

$$u'(C(k)) = \beta u'(C(F(k - C(k)))) F'(k - C(k)).$$

Standard linearization methods produce $C'(k)$. Successive differentiations of (10) produce higher-order derivatives of $C(k)$ at $k = k^*$. For example, the second derivative of (10) together with the steady-state condition $k = k^*$ implies that $C''(k^*)$ satisfies the linear equation

$$\begin{aligned} u''C'' + u'''C'C' = & \beta u'''(C'F'(1 - C'))^2 F' + \beta u''C''(F'(1 - C'))^2 F' \\ & + 2\beta u''C'F'(1 - C')^2 F'' + \beta u'F'''(1 - C')^2 \\ & + \beta u'F''(-C'') \end{aligned}$$

where the parentheses denote multiplication, not application of a function. All functions are evaluated at the steady state value of their arguments. This is a linear equation in the unknown $C''(k^*)$. Linear operations combined with successive differentiations of (10) produce all higher-order derivatives.

The solution in the general case is a policy function, $C(k, \varepsilon)$, which expresses consumption as a function of the capital stock k as well as the standard deviation ε . $C(k, \varepsilon)$ satisfies the Euler equation

$$\begin{aligned} u'(C(k, \varepsilon)) &= \beta E \{u'(g(\varepsilon, k, z)) R(\varepsilon, k, z)\} \\ g(\varepsilon, k, z) &\equiv C((1 + \varepsilon z)F(k - C(k))) \\ R(\varepsilon, k, z) &\equiv (1 + \varepsilon z)F'(k - C(k)) \end{aligned} \quad (11)$$

Differentiation of (11) with respect to ε produces a linear equation for $C_\varepsilon(k^*, 0)$, which has the solution $C_\varepsilon = 0$. This is a natural result since ε parameterizes the standard deviation of uncertainty, whereas basic economic intuition says that the economic response should be proportional to variance, which is ε^2 . Furthermore, the perturbation variable in Fleming (1971) was instantaneous variance. Therefore, $C_\varepsilon = 0$ is a natural result.

Further differentiation with respect to ε produces a linear equation for $C_{\varepsilon\varepsilon}(k^*, 0)$ shows that

$$C_{\varepsilon\varepsilon}(k^*, 0) = \frac{u''' C' C' F^2 + 2u'' C' F + u'' C'' F^2}{u'' C' F' + \beta u' F''}$$

where all the derivatives of u and F are evaluated at the steady-state values of c and k . This can be continued to compute higher-order derivatives as long as u and F have the necessary derivatives.

It may initially appear more natural to use variance, ε^2 , as the perturbation variable since $C_\varepsilon(k, 0) = 0$. However, using the variance would cause difficulty in a discrete-time problem. The ε^3 term $C_{\varepsilon\varepsilon\varepsilon}$ is nonzero in the deterministic case since skewness can be nonzero. This is not a problem in the continuous-time, Ito process case since all odd (instantaneous) moments are zero. A Taylor series in ε^2 in discrete-time stochastic problems would miss the ε^3 terms and would fail at or before the ε^3 term. In terms of asymptotic theory, the appropriate gauge for discrete-time stochastic problems is ε^k instead of ε^{2k} .

4. PERTURBATIONS OF THE DETERMINISTIC MODEL

We now begin applying perturbation methods to rational expectations models of the form in (1). We first describe the perturbation method for the deterministic problem. The deterministic problem has independent interest. Since the perturbations are with respect to the state variables x , we drop the ε parameter in this section to simplify the notation.

Suppose that $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. Then $Y(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}^m$, and $F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$. Note that we assume that the number of equations in $g = 0$ equals the number of free variables, m . We express equilibrium in the more convenient form

$$0 = \mathcal{G}(x, Y(x), \mathcal{X}(x), \mathcal{Y}(x)) \equiv \begin{pmatrix} g^\mu(x, Y(x), \mathcal{X}(x), \mathcal{Y}(x)) \\ \mathcal{X}^I(x) - F^I(x, Y(x)) \\ \mathcal{Y}^A(x) - Y^A(\mathcal{X}(x)) \end{pmatrix} \quad (12)$$

where g^μ denotes the μ 'th equation in the collection of equilibrium equations. This formulation uses expressions $Y(x)$ and $F(x, Y(x))$ as well as the composite expressions $\mathcal{X}(x) - F(x, Y(x))$ and $\mathcal{Y}(x) = Y(\mathcal{X}(x)) = Y(F(x, Y(x)))$. The introduction of the intermediate terms $\mathcal{X}(x)$ and $\mathcal{Y}(x)$ helps us make clear the essentially linear structure of the problem. It also indicates a direction for efficient programming since it tells us that we should separately compute the derivatives of $\mathcal{X}(x)$ and $\mathcal{Y}(x)$ before we compute the derivatives of $g(x, Y(x), \mathcal{X}(x), \mathcal{Y}(x))$. This approach also distinguishes the cases where Y occurs as $Y(x)$ as opposed to $Y(F(x, Y(x)))$. Y^α will refer to occurrences of $Y(x)$ and Y^A will refer to occurrences of $Y(F(x, Y(x)))$. We used F^I , F^J , etc., to refer to components of $F(x, Y(x)) = \mathcal{X}(x)$. \mathcal{X}^I , \mathcal{X}^J , etc., will also refer to components of $\mathcal{X}(x)$, and components of $\mathcal{Y}(x) = Y(F(x, Y(x)))$ will be denoted \mathcal{Y}^A , \mathcal{Y}^B , \mathcal{Y}^Δ , etc.

The objective is to find the derivatives of $Y(x)$ with respect to x at the deterministic steady state, and use that information to construct Taylor series approximations of $Y(x)$. In conventional notation, that Taylor series is expressed as

$$Y(x) \doteq y_* + Y_x(x_*, 0)(x - x_*) + (x - x_*)^\top Y_{xx}(x_*, 0)(x - x_*) + \dots$$

but tensor notation expresses it as

$$\begin{aligned} Y(x) \doteq & y_* + Y_i(x_*, 0)(x^i - x_*^i) + Y_{ij}(x_*, 0)(x^i - x_*^i)(x^j - x_*^j) \\ & + Y_{ijk}(x_*, 0)(x^i - x_*^i)(x^j - x_*^j)(x^k - x_*^k) + \dots \end{aligned}$$

Deterministic Steady State and the Linear Approximation. Perturbation methods begin with the deterministic steady state, which is the solution to

$$\begin{aligned} 0 &= g(x_*, y_*, x_*, y_*) \\ x_* &= F(x_*, y_*) \end{aligned}$$

This is a nonlinear set of equations. The second step in perturbation methods is to compute the linear terms of the approximation $Y_x(x_*, 0)$. Standard linearization methods show that the coefficients $Y_x(x_*, 0)$ are the solution, $y = y_* + Y_x(x - x_*)$, to the linear rational expectations model

$$g_i^\mu x_t^i + g_\alpha^\mu y_t^\alpha + g_I^\mu x_{t+1}^I + g_A^\mu y_{t+1}^A = 0 \quad (13)$$

where all the gradients of g^μ in (13) are evaluated at the deterministic steady state. We assume locally unique and stable dynamics. This may not be true for all steady states; we confine our attention to only saddlepoint stable steady states. This linearization procedure is justified by the implicit function theorem applied to (6). The solution is

$$y_t - y_* = H(x_t - x_*) \quad (14)$$

Anderson et al. (1996) and Anderson, Evans, Sargent, and McGrattan (1996) survey methods for solving such models. This is the difficult step computationally, but can be handled by conventional methods.

Theorem 1. *If g and F are locally analytic in a neighborhood of (x_*, y_*, x_*, y_*) , and (13) has a unique locally asymptotically stable solution, then for some $\varepsilon > 0$ there is a unique function $Y(x)$ such that $Y(x)$ solves (12) and is locally analytic for $\|x - x_*\| < \varepsilon$. In particular, $Y(x)$ is infinitely differentiable and its derivatives solve the equations derived by implicit differentiation (12).*

Proof. This follows from the standard application of the analytic implicit function theorem to the space of sequences that converge to the steady state at an exponential rate. ■

Before we move on to the second-order approximation, we need to formulate the first order problem in functional and tensor form. We start with equilibrium expressed in the form (12). Let $\mathcal{Z}(x) = (x, Y(x), \mathcal{X}(x), \mathcal{Y}(x))$. The first-order derivatives with respect to the x^i variables are

$$\begin{aligned} 0 &= g_i^\mu(\mathcal{Z}(x)) + g_\alpha^\mu(\mathcal{Z}(x)) Y_i^\alpha(x) + g_I^\mu(\mathcal{Z}(x)) \mathcal{X}_i^I(x) + g_A^\mu(\mathcal{Z}(x)) \mathcal{Y}_i^A(x) \\ \mathcal{X}_i^I(x) &= F_i^I(x, Y(x)) + F_\alpha^I(x, Y(x)) Y_i^\alpha \\ \mathcal{Y}_i^A(x) &= Y_I^A(\mathcal{X}(x)) \mathcal{X}_i^I(x) \end{aligned} \tag{15}$$

At the steady state (we drop arguments here since they are all understood to have their steady state values)

$$\begin{aligned} 0 &= g_i^\mu + g_\alpha^\mu Y_i^\alpha + g_I^\mu \mathcal{X}_i^I + g_A^\mu \mathcal{Y}_i^A \\ \mathcal{X}_i^I &= F_i^I + F_\alpha^I Y_i^\alpha \\ \mathcal{Y}_i^A &= Y_I^A \mathcal{X}_i^I \end{aligned}$$

After substitution, we see that the tensor (matrix) Y_i^α is the solution to

$$0 = g_i^\mu + g_\alpha^\mu Y_i^\alpha + g_I^\mu (F_i^I + F_\alpha^I Y_i^\alpha) + g_A^\mu Y_I^A (F_i^I + F_\alpha^I Y_i^\alpha)$$

This is a matrix polynomial equation for the matrix $Y_i^\alpha(x_*)$ with n^2 equations in the n^2 unknown values Y_i^α . This is also the H matrix in (14) computed by standard methods.

Before we move to higher order approximations, we express the first order system in (15) in a convenient form. Let

$$\mathcal{G}(\mathcal{Z}(x))_{\mathbb{N}} = \begin{pmatrix} g_{\alpha}^{\mu}(\mathcal{Z}(x)) & g_I^{\mu}(\mathcal{Z}(x)) & g_A^{\mu}(\mathcal{Z}(x)) \\ -F_{\alpha}^I(x, Y(x)) & \mathbf{1}_n & 0 \\ 0 & 0 & \mathbf{1}_m \end{pmatrix}$$

The (α, I, A) subscript notation represents the fact that $\mathcal{G}(\mathcal{Z}(x))_{\mathbb{N}}$ is three columns of tensors, the first being derivatives with respect to α , etc. The $\mathbf{1}_n$ ($\mathbf{1}_m$) entry in $\mathcal{G}(\mathcal{Z}(x))_{\mathbb{N}}$ represents the $n \times n$ ($m \times m$) identity map². Then the system of equations in (15) defining the first-order coefficients can be expressed as

$$0 = \mathcal{G}(\mathcal{Z}(x))_{\mathbb{N}} \begin{pmatrix} Y_i^{\alpha}(x) \\ \mathcal{X}_i^I(x) \\ \mathcal{Y}_i^A(x) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ Y_I^A(\mathcal{X}(x))\mathcal{X}_i^I(x) \end{pmatrix} + \begin{pmatrix} g_i^{\mu}(\mathcal{Z}(x)) \\ F_i^I(x, Y(x)) \\ 0 \end{pmatrix} \quad (16)$$

and the steady state values $Y_i^{\alpha}(x_*)$ satisfy

$$0 = \mathcal{G}(z_*)_{\mathbb{N}} \begin{pmatrix} Y_i^{\alpha}(x_*) \\ \mathcal{X}_i^I(x_*) \\ \mathcal{Y}_i^A(x_*) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ Y_J^A(x_*)\mathcal{X}_i^J(x_*) \end{pmatrix} + \begin{pmatrix} g_i^{\mu}(z_*) \\ F_i^I(x_*, y_*) \\ 0 \end{pmatrix}$$

where $z_*^* = \mathcal{Z}(x_*)$. We will use the form in (16) below.

Second-order Approximation. We next want to compute $Y_{ij}^{\alpha}(x_*)$, the Hessian of $Y^{\alpha}(x)$ at $x = x_*$. Let $D_j f(x)$ represent the total derivative of $f(x)$ w.r.t. x^j . Dif-

²We use the $\mathbf{1}_n$ notation instead of the Kronecker delta tensor δ_j^i notation since we do not want to change indices.

ferentiation of (16) with respect to x^j shows

$$\begin{aligned}
0 = & D_j (\mathcal{G}(\mathcal{Z}(x))_{\mathbb{N}}) \begin{pmatrix} Y_i^\alpha(x) \\ \mathcal{X}_i^I(x) \\ \mathcal{Y}_i^A(x) \end{pmatrix} + \mathcal{G}(\mathcal{Z}(x))_{(\alpha, I, A)} \begin{pmatrix} Y_{ij}^\alpha(x) \\ \mathcal{X}_{ij}^I(x) \\ \mathcal{Y}_{ij}^A(x) \end{pmatrix} \\
& - \begin{pmatrix} 0 \\ 0 \\ Y_{IJ}^A(\mathcal{X}(x)) \mathcal{X}_j^J(x) \mathcal{X}_i^I(x) + Y_I^A(\mathcal{X}(x)) \mathcal{X}_{ij}^I(x) \end{pmatrix} + D_j \begin{pmatrix} g_i^\mu(\mathcal{Z}(x)) \\ F_i^I(x, Y(x)) \\ 0 \end{pmatrix}
\end{aligned} \tag{17}$$

which implies the steady state conditions

$$\begin{aligned}
0 = & D_i (\mathcal{G}(z_*)) \begin{pmatrix} Y_i^\alpha(x_*) \\ \mathcal{X}_i^I(x_*) \\ \mathcal{Y}_i^A(x_*) \end{pmatrix} + D_i \begin{pmatrix} g_i^\mu(z_*) \\ F_i^I(x_*, Y(x_*)) \\ 0 \end{pmatrix} \\
& - \begin{pmatrix} 0 & 0 & 0 \\ \mathcal{X}_j^J(x_*) \mathcal{X}_i^I(x_*) & Y_I^A(x_*) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} Y_{IJ}^A(x_*) \\ \mathcal{X}_{ij}^I(x_*) \\ \mathcal{Y}_{ij}^A(x_*) \end{pmatrix} \\
& + \mathcal{G}(z_*)_{\mathbb{N}} \begin{pmatrix} Y_{ij}^\alpha(x_*) \\ \mathcal{X}_{ij}^I(x_*) \\ \mathcal{Y}_{ij}^A(x_*) \end{pmatrix}
\end{aligned} \tag{18}$$

which is a linear equation in the unknowns $(Y_{ij}^\alpha(x_*), \mathcal{X}_{ij}^I(x_*), \mathcal{Y}_{ij}^A(x_*))$. Note here that the solvability condition is the nonsingularity of $\mathcal{G}(z_*)$ plus some other terms.

Theorem 2. $(Y_{ij}^\alpha(x_*), \mathcal{X}_{ij}^I(x_*), \mathcal{Y}_{ij}^A(x_*))$ satisfies the linear system (18). It is uniquely solved by (18) if and only if (18) is nonsingular

Third-order approximation. The third-order terms $\mathcal{Z}_{ijk}^{\aleph}(x)$ are found by differentiating (17) with respect to x^j , producing

$$\begin{aligned}
0 = & D_k \left(D_j (\mathcal{G}(\mathcal{Z}(x))_{\aleph}) \begin{pmatrix} Y_i^{\alpha}(x) \\ \mathcal{X}_i^I(x) \\ \mathcal{Y}_i^A(x) \end{pmatrix} + D_j \begin{pmatrix} g_i^{\mu}(\mathcal{Z}(x)) \\ F_i^I(x, Y(x)) \\ 0 \end{pmatrix} \right) \\
& + D_k (\mathcal{G}(\mathcal{Z}(x))_{\aleph}) \begin{pmatrix} Y_{ij}^{\alpha}(x) \\ \mathcal{X}_{ij}^I(x) \\ \mathcal{Y}_{ij}^A(x) \end{pmatrix} + \mathcal{G}(\mathcal{Z}(x))_{\aleph} \begin{pmatrix} Y_{ijk}^{\alpha}(x) \\ \mathcal{X}_{ijk}^I(x) \\ \mathcal{Y}_{ijk}^A(x) \end{pmatrix} \\
& - \begin{pmatrix} 0 \\ D_{ijk} Y^A(\mathcal{X}(x)) \\ 0 \end{pmatrix}
\end{aligned}$$

The steady state is now given by

$$0 = (\text{Terms without } \mathcal{Z}_{ijk}) + \mathcal{G}(\mathcal{Z}(x))_{\aleph} \begin{pmatrix} Y_{ijk}^{\alpha}(x_*) \\ \mathcal{X}_{ijk}^I(x_*) \\ \mathcal{Y}_{ijk}^A(x_*) \end{pmatrix} - \begin{pmatrix} 0 \\ D_{ijk} Y^A(\mathcal{X}(x_*)) \\ 0 \end{pmatrix} \quad (19)$$

where

$$\begin{aligned}
D_{ijk} Y^A(\mathcal{X}(x_*)) &= Y_{IJK}^A(x_*) \mathcal{X}_k^K(x_*) \mathcal{X}_j^J(x_*) \mathcal{X}_i^I(x_*) \\
&+ Y_{IJ}^A(x_*) D_{x_k} (\mathcal{X}_j^J(x_*) \mathcal{X}_i^I(x_*)) \\
&+ Y_{IK}^A(x_*) \mathcal{X}_k^K(x_*) \mathcal{X}_{ij}^I(x_*) + Y_I^A(x_*) \mathcal{X}_{ijk}^I(x_*)
\end{aligned}$$

Theorem 3. $(Y_{ijk}^{\alpha}(x_*), \mathcal{X}_{ijk}^I(x_*), \mathcal{Y}_{ijk}^A(x_*))$ satisfies the linear system (19). It is uniquely solved by (19) if and only if (19) is nonsingular

There are two items to note. First, $Y_{ijk}^{\alpha}(x_*)$ satisfies a linear system of equations. Second, the solvability matrix for $\mathcal{Z}_{ijk}(x_*)$ will be different than the solvability matrix for $\mathcal{Z}_{ij}(x_*)$. The fact that the solvability conditions change as we change order

is a potential problem. However, one suspects that these matrices are generically determinate, but that remains an open issue.

The following is an obvious continuation of our method.

Theorem 4. *Given the solution to all lower order derivatives, the degree m derivatives $(Y_{i_1 \dots i_m}^\alpha(x_*), \mathcal{X}_{i_1 \dots i_m}^I(x_*), \mathcal{Y}_{i_1 \dots i_m}^A(x_*))$ satisfy a linear equation that is solvable if the linear map*

$$\mathcal{G}(z_*)_{\mathbb{R}} \begin{pmatrix} Y_{i_1 \dots i_m}^\alpha(x_*) \\ \mathcal{X}_{i_1 \dots i_m}^I(x_*) \\ \mathcal{Y}_{i_1 \dots i_m}^A(x_*) \end{pmatrix} - \begin{pmatrix} 0 \\ Y_{I_1 \dots I_m}^A(x_*) \mathcal{X}_{i_1}^{I_1}(x_*) \cdots \mathcal{X}_{i_m}^{I_m}(x_*) + Y_I^A(x_*) \mathcal{X}_{i_1 \dots i_m}^I(x_*) \\ 0 \end{pmatrix}$$

is invertible.

4.1. Algorithm. We have established that Taylor expansions of the equilibrium equations produce a series of linear equations in the derivatives of Y once we get past the first-order terms. This implies that a fairly simple algorithm can be applied once we have the linear systems. Define

$$G^\mu(x) = g^\mu(x, Y(x), \mathcal{X}(x), \mathcal{Y}(x))$$

where

$$\begin{aligned} \mathcal{X}^I(x) &= F^I(x, Y(x)) \\ \mathcal{Y}^A(x) &= Y^A(\mathcal{X}(x)) \end{aligned}$$

The steady state definition tells us that $G(x_*) = 0$. Furthermore, if $Y_i^\alpha(x_*)$ is fixed at its true linear solution, $G_i^\mu(x_*) = 0$. The preceding calculations show that the

Taylor expansion continues with the form

$$\begin{aligned}
G^\mu(x_* + \varepsilon v) &= G^\mu(x_*) + \varepsilon G_i^\mu(x_*) v^i \\
&+ \varepsilon^2 (M_\alpha^{\mu,2}(x_*) + N_\alpha^{\mu,2}(x_*) Y_{ij}^\alpha(x_*)) v^i v^j \\
&+ \varepsilon^3 (M_\alpha^{\mu,3}(x_*) + N_\alpha^{\mu,3}(x_*) Y_{ijk}^\alpha(x_*)) v^i v^j v^k \\
&+ \dots
\end{aligned} \tag{20}$$

where each $M_\alpha^{\mu,\ell}(x_*)$ and $N_\alpha^{\mu,\ell}(x_*)$ term involves no derivative of $Y^\alpha(x)$ of order ℓ or higher. Therefore, we take a specific problem, have the computer produce the Taylor series expansion in (20) where the $Y_i^\alpha(x_*)$, $Y_{ij}^\alpha(x_*)$, etc., terms are left free, and then compute them in a sequentially linear fashion by solving in sequence the linear systems

$$\begin{aligned}
0 &= M_\alpha^{\mu,2}(x_*) + N_\alpha^{\mu,2}(x_*) Y_{ij}^\alpha(x_*) \\
0 &= M_\alpha^{\mu,3}(x_*) + N_\alpha^{\mu,3}(x_*) Y_{ijk}^\alpha(x_*) \\
&\dots
\end{aligned}$$

where at each stage we use the solutions in the previous stage.

5. A GENERAL STOCHASTIC PROBLEM

We next compute the stochastic portion of our approximation. The stochastic rational expectations problem is

$$\begin{aligned}
0 &= E \{ g^\mu(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z) | x_t \}, \quad \mu = 1, \dots, m \\
x_{t+1}^i &= F^i(x_t, y_t, \varepsilon z_t), \quad i = 1, \dots, n
\end{aligned}$$

The objective is to find some equilibrium rule, $Y(x, \varepsilon)$, such that

$$E \{ g(x, Y(x, \varepsilon), F(x, Y(x, \varepsilon), \varepsilon z), Y(F(x, Y(x, \varepsilon), \varepsilon z), \varepsilon), \varepsilon z) | x \} = 0 \tag{21}$$

We have constructed the derivatives of $Y(x, 0)$ with respect to the components of x . We now compute the ε derivatives.

Before we describe our method, we first present an example which highlights the pitfalls of pursuing a Taylor series expansion procedure in a casual fashion ignoring the relevant mathematics. We will posit a simple example of a rational expectations model and follow the approach taken in Kydland and Prescott (1982), and elsewhere. This casual application of the standard approach will produce a nonsensical result and highlight a potential problem. We will then proceed to develop an approach consistent with the implicit function theorem.

5.1. A Cautionary Example. We now present a simple example which highlights the dangers of a casual approach to computing Taylor series expansions. Suppose that an investor receives K endowment of wealth and that there are two possible investments, I_1 and I_2 ; therefore, $I_{1,t} + I_{2,t} = K$. We allow negative values for $I_{1,t}$ and $I_{2,t}$, making this example like a portfolio problem. Assume a gross “adjustment cost” for deviations of type 1 investment from some target \bar{I} equal to $\alpha (I_{1,t} - \bar{I})^2 / 2$ units of utility. In period $t+1$ the investments produce $f(I_{1,t}, I_{2,t}, Z_{t+1})$ of the consumption good. Assume $f(I_{1,t}, I_{2,t}, Z_{t+1}) = I_{1,t} + I_{2,t}Z_{t+1}$ where Z_t is log Normal with mean $1 + \mu$ and variance σ^2 . Assume that all of the second-period gross return is consumed and that the utility function is $u(c) = c^{1-\gamma} / (1 - \gamma)$. The complete utility function is

$$U(I_{1,t}, I_{2,t}, Z_{t+1}) = -\alpha (I_{1,t} - \bar{I})^2 / 2 + E_t \{u(I_{1,t} + I_{2,t}Z_t)\} \quad (22)$$

This is a rational expectations model where the endogenous variables are defined by the equations

$$\begin{aligned} 0 &= -\alpha (I_{1,t} - \bar{I}) + E_t \{u'(I_{1,t} + I_{2,t}Z_t) Z_t\} \\ \bar{K} &= I_{1,t} + I_{2,t} \end{aligned} \quad (23)$$

This is a trivial rational expectations model³, but if a method cannot reliably approximate the solution to this problem then we would not have any confidence in its general validity.

The deterministic “steady state” is $I_{1,t} = \bar{I}$, $I_{2,t} = K - \bar{I}$; that is, we set type 1 investment equal to the type 1 target \bar{I} and put the rest of the capital in type 2 investments.⁴ This is obvious since the investments have the same return and there is an adjustment cost for any deviations of $I_{1,t}$ from zero. We now want to know how the investment policy is altered if we add some noise to the risky type 2 investment. If we were to take a certainty equivalent approximation then the investment rules are unchanged by an increase in variance. However, this will produce nonsensical answers if $\bar{I} < 0$ and $I_{2,t} > K$ since this would imply that there is some chance that $I_{1,t} + I_{2,t}Z_t < 0$, implying negative consumption. A more sophisticated application of the key idea in Kydland and Prescott (1982)⁵ tells us to replace the utility function in (22) with a quadratic approximation around the deterministic consumption K and solve the resulting linear equation for $I_{1,t}$. For any \bar{I} , μ , and σ^2 , as α goes to zero we converge to the limiting investment rule

$$I_{2,t} = K \frac{\mu/\gamma}{\sigma^2 + \mu^2} \quad (24)$$

Consider the situation when $\mu/\gamma > \sigma^2 + \mu^2$, as is the case when $\mu = .06$ and $\sigma^2 = .04$, the standard calibration for equity investment, and $\gamma = 1$, which is log utility. In this case $I_{1,t}$, the safe investment, would be negative, implying that one borrows

³This example is also a bit silly with the utility adjustment costs, but it is an attempt to construct an example which looks like a portfolio problem (think of I_1 and I_2 as investments in alternative securities) but avoids the complications examined in Judd and Guu (2001).

⁴Note that if $\bar{I} < 0$ then $I_{2,t} > K$.

⁵Strictly speaking, Kydland and Prescott assume additive noise in the law of motion for the state, in which case their approximation is certainty equivalent. However, the application of their key linear-quadratic approximation idea to this example with nonadditive disturbances is clear.

($I_{1,t} < 0$) to buy more than K dollars of the risky investment. This, however, is nonsensical since the support of log Normal Z is all of the nonnegative real line, implying that there is some chance of negative consumption, which is inconsistent with CRRA utility. The possibility of a negative consumption makes expected utility undefined and can be avoided⁶ by choosing any nonnegative $I_{1,t}$. Therefore, the ad hoc approximation in (24) cannot be a useful approximation to the the solution of (23).

What went wrong? There is an implicit constraint on consumption being positive in (23) since the utility function is not defined for negative consumption. However, that constraint is not binding in the deterministic problem since consumption is surely positive. Since this constraint is not present in the deterministic case, it is not present when one executes a purely local procedure. If there are no restrictions placed on the random disturbances, the Kydland and Prescott (1982) procedure implicitly replaces a nonlinear problem with a linear-quadratic problem *globally*. The approximate quadratic utility function is defined for negative consumption and is not an appropriate *global* approximation since the true utility function may not be defined for negative consumption. Therefore, without restrictions on the disturbances, the Kydland and Prescott approach is not a local analysis based on some implicit function theorem.

This is not a problem unique to the strategy recommended by Kydland and Prescott. The second-order procedure of Sims (2000) is a natural extension of Kydland and Prescott and would fail in this case for similar reasons. In fact, any scheme using purely local information and moments and ignoring global considerations can fail on problems like this since local information cannot alone model global considerations. The approximation procedure advocated in Campbell (2002) is different

⁶The reason for choosing log Normal z is obvious; if we had made z Normal than no nonzero choice for I_2 is consistent with the nonnegativity constraint on consumption.

but also fails on this point, often producing approximations with positive probabilities of negative consumption. Campbell's approach begins with the observation that one can analytically compute expected utility if utility is CRRA and consumption is log Normal. Therefore, it would be nice if our portfolio problems always reduced to computing the expected CRRA utility of log Normal consumption. However, if asset returns are log Normal, consumption will generally not be log Normal since nontrivial linear combinations of log Normal returns are not log Normal. Campbell approximates the original portfolio problem (our problem in (23) is a portfolio problem if $\alpha = 0$) by replacing the ex post distribution of consumption with an "approximating" log Normal distribution, and then finds that portfolio which maximizes the approximate expression for expected utility. Of course, the support of the log Normal approximation will not include any negative values and if mean returns are sufficiently large relative to the variance and risk aversion, the Campbell "approximation" will short bonds and cause a positive probability of negative consumption.

These problems do not go away even if we take the time period to zero. If we were to embed (23) in a sequence of problems with ever shorter time periods, we would want to maintain the Sharpe ratio, μ/σ^2 , above some positive limit. Then, for many γ , the approximation in (24) would always be greater than K , implying shorting in each discrete-time problem and a positive probability of negative consumption.

The mathematical economics literature has long been aware of this problem and has recognized the importance of proceeding locally. In particular, Samuelson (1970) noted this problem and assumed that the disturbance Z has bounded support. Judd and Guu (2001) generalized this to a local analyticity condition. Judd (1998) (see chapter 16) also points out the importance of controlling the distribution of Z . In this paper, we will proceed with a bounded support assumption since that is the most general way to proceed. We will display the critical conditions which need to be checked before one can proceed with a perturbation approximation of rational

expectations models.

5.2. Local Perturbation Approach for Stochastic Models. We have already computed the Taylor series approximation of the deterministic problem

$$\begin{aligned} Y^\alpha(x, \varepsilon) &\doteq y_* + Y_i^\alpha(x_*, 0) (x^i - x_*^i) \\ &\quad + Y_{ij}^\alpha(x_*, 0) (x^i - x_*^i) (x^j - x_*^j) \\ &\quad + Y_{ijk}^\alpha(x_*, 0) (x^i - x_*^i) (x^j - x_*^j) (x^k - x_*^k) \\ &\quad + \dots \end{aligned}$$

We next move to the stochastic terms, $Y_\varepsilon(x_*, 0)$, $Y_{\varepsilon\varepsilon}(x_*, 0)$, etc. The general stochastic system is

$$\begin{aligned} 0 &= E \{ g(x_t, y_t, x_{t+1}, y_{t+1}, \varepsilon z) | x_t \} \\ x_{t+1} &= F(x_t, y_t, \varepsilon z_t) \end{aligned}$$

which implies the functional equation

$$\begin{aligned} 0 &= E \{ g^\mu(x, Y(x, \varepsilon), F(x, Y(x, \varepsilon), \varepsilon z), Y(F(x, Y(x, \varepsilon), \varepsilon z)), \varepsilon z) \} \\ &= \mathcal{N}(Y)(x, \varepsilon) \end{aligned} \quad (25)$$

holds at all (x, ε) . Differentiation with respect to ε shows

$$\begin{aligned} 0 &= E \{ g_\alpha^\mu Y_\varepsilon^\alpha + g_I^\mu (F_\alpha^I Y_\varepsilon^\alpha + F_\varepsilon^I z) + g_A^\mu (Y_I^A (F_\alpha^I Y_\varepsilon^\alpha + F_\varepsilon^I z) + Y_\varepsilon^A) + z g_\varepsilon^\mu \} \\ &= E \{ g_\alpha^\mu Y_\varepsilon^\alpha \} + E \{ g_I^\mu F_\alpha^I Y_\varepsilon^\alpha \} + E \{ g_I^\mu F_\varepsilon^I z \} + E \{ g_A^\mu Y_I^A F_\alpha^I Y_\varepsilon^\alpha \} \\ &\quad + E \{ g_A^\mu Y_I^A F_\alpha^I F_\varepsilon^I z \} + E \{ g_A^\mu Y_\varepsilon^A \} + E \{ z g_\varepsilon^\mu \} \\ &= Y_\varepsilon^\alpha (E \{ g_\alpha^\mu \} + E \{ g_I^\mu F_\alpha^I \} + E \{ g_A^\mu Y_I^A F_\alpha^I \} + \delta_\alpha^A E \{ g_A^\mu \}) \\ &\quad + E \{ g_I^\mu F_\varepsilon^I z \} + E \{ g_A^\mu Y_I^A F_\alpha^I F_\varepsilon^I z \} + E \{ z g_\varepsilon^\mu \} \end{aligned} \quad (26)$$

holds at all $(x, \varepsilon)^7$. Note that in the last step we use the identity $Y_\varepsilon^A = \delta_\alpha^A Y_\varepsilon^\alpha$. At the steady state of the deterministic case, $(x, \varepsilon) = (x_*, 0)$, the εz terms collapse to zero, the derivatives g_α^μ , g_I^μ , g_A^μ , F_α^I , F_ε^I , and g_ε^μ become deterministic, and (26) reduces to

$$0 = Y_\varepsilon^\alpha (g_\alpha^\mu + g_I^\mu F_\alpha^I + g_A^\mu Y_I^A F_\alpha^I + \delta_\alpha^A g_A^\mu) + (g_\varepsilon^\mu + g_I^\mu F_\varepsilon^I + g_A^\mu Y_I^A F_\alpha^I F_\varepsilon^I) E\{z\} \quad (27)$$

which has a unique solution for Y_ε^α iff all the terms in (26) exists at $\varepsilon = 0$ and

$$\mathcal{N}_Y = g_\alpha^\mu + g_I^\mu F_\alpha^I + g_A^\mu Y_I^A F_\alpha^I + \delta_\alpha^A g_A^\mu$$

is an invertible matrix. The problem of existence arises with the terms g_ε^μ and F_ε^I which only arise here. As we saw above, there are examples where these terms do not exist.

We can take higher-order derivatives of (25) with respect to ε to arrive at equations for $Y_{\varepsilon\varepsilon}^\alpha$, $Y_{\varepsilon\varepsilon}^\alpha$, etc. The formulas are complex, however, the pattern is clear. For example, the derivative of (26) with respect to ε implies

$$\begin{aligned} 0 &= Y_{\varepsilon\varepsilon}^\alpha (E\{g_\alpha^\mu\} + E\{g_I^\mu F_\alpha^I\} + E\{g_A^\mu Y_I^A F_\alpha^I\} + \delta_\alpha^A E\{g_A^\mu\}) \\ &\quad Y_\varepsilon^\alpha \frac{d}{d\varepsilon} (E\{g_\alpha^\mu\} + E\{g_I^\mu F_\alpha^I\} + E\{g_A^\mu Y_I^A F_\alpha^I\} + \delta_\alpha^A E\{g_A^\mu\}) \\ &\quad + \frac{d}{d\varepsilon} (E\{g_I^\mu F_\varepsilon^I z\} + E\{g_A^\mu Y_I^A F_\alpha^I F_\varepsilon^I z\} + E\{z g_\varepsilon^\mu\}) \end{aligned}$$

which in turn implies that $Y_{\varepsilon\varepsilon}^\alpha(x_*, 0)$ solve an equation of the form

$$0 = Y_{\varepsilon\varepsilon}^\alpha \mathcal{N}_Y + \mathcal{M}$$

where \mathcal{M} contains only derivatives of g and F and moments of z . This will determine $Y_{\varepsilon\varepsilon}^\alpha(x_*, 0)$ if the terms in \mathcal{M} exists and \mathcal{N}_Y is invertible. Notice that the

⁷We are slightly abusing notation by writing

$$\frac{d}{d\varepsilon} g(x, y, \hat{x}, \hat{y}, \varepsilon z) = g_\varepsilon(x, y, \hat{x}, \hat{y}, \varepsilon z) z$$

More formally, this should be $g_5(x, y, \hat{x}, \hat{y}, \varepsilon z) z$ but we use the $g_\varepsilon(\dots)$ notation for its mnemonic advantage. The same comment applies to our notation $F_\varepsilon(\dots)$.

solvability condition, the invertibility of \mathcal{N}_Y , is the same as the solvability condition for $Y_\varepsilon^\alpha(x_*, 0)$. Continuing this process shows that the ε derivatives of Y exist as long as g and F have the necessary derivatives and the moments of z exist. The next theorem summarizes our argument under the assumption that g is analytic.

Theorem 5. *If (i) $g^\mu(x, y, \hat{x}, \hat{y}, \varepsilon z)$ exists and is analytic for all z in some neighborhood of $(x, y, \hat{x}, \hat{y}, \varepsilon) = (x_*, y_*, x_*, y_*, 0)$, (ii) there exists a unique deterministic solution $Y(x, 0)$ locally analytic in x , and (iii) \mathcal{N}_Y is invertible, then there is an $r > 0$ such that for all $(x, \varepsilon) \in B_r(x_*, 0)$, there exists a unique solution $Y(x, \varepsilon)$ to (25). Furthermore, all derivatives of $Y(x, \varepsilon)$ exist in a neighborhood of $(x_*, 0)$ and can be solved by implicit differentiation.*

Proof. Application of the implicit function theorem for analytic functions. See Zeidler (1986). ■

This theorem sounds obvious, but the conditions are important ones and should be checked. In particular, our example in (23) fails to satisfy condition (i). One natural and general way to insure (i) is to assume z has bounded support. This may sound limiting since many popular processes in dynamic analyses, such as linear processes with Normal or log Normal. Also, continuous time Ito processes often imply finite-horizon distributions with infinite support. However, assuming z has bounded support is a suitable assumption for discrete-time models, and, in fact, may be a superior way to approximate continuous-time problems. Consider, for example, Merton's (1972) portfolio analysis. For some parameters, investors will short the bond market and use the proceeds to buy stocks even for log utility. Our example above shows that this would be unwise in a discrete-time model where returns have log Normal returns since there would be some chance that wealth goes negative. Therefore, assuming log Normal returns in a discrete-time model is a poor way to approximate the continuous-time model. Why is there such a difference? In the

continuous-time model with an Ito process driving the returns, an investor hit with a series of negative return shocks can reduce his exposure to risk before his wealth goes negative, but this is impossible in the discrete-time model and log Normal returns. The finite-support assumption for z similarly allows an investor facing a series of negative shocks to adjust his position before hitting ruin.

The key point is that there are regularity conditions which must be satisfied if we are to use the implicit function theorem to justify asymptotic expansions as approximations. Our example shows that it may not be possible to combine popular utility functions with popular stochastic processes. We take the view that it is more important to accurately approximate the economic process than to stay with popular stochastic processes. Therefore, we assume Assumption 1:

Assumption 1: *The support of z is finite and $E\{z\} = 0$.*

The following theorem follows directly from $E\{z\} = 0$ and successive differentiation of (25). It just ratifies common sense that a first-order change in standard deviation affects nothing, and that the dominant order effect is variance, which here equals ε^2 .

Theorem 6. *If $E\{z\} = 0$, then $0 = Y_\varepsilon^\alpha = Y_{\varepsilon i}^\alpha = Y_{\varepsilon ij}^\alpha = \dots$*

Proof. $0 = Y_\varepsilon^\alpha$ follows from (27) and $E\{z\} = 0$. The other results follow from the fact that derivatives of (26) with respect to state variables in x will always reduce to expressions of the form $0 = Y_{\varepsilon i_1 \dots i_\ell}^\alpha(\dots) + (\dots) E\{z\}$. ■

6. NONLOCAL ACCURACY TESTS

Perturbation methods produce the best possible asymptotically valid local approximations to a problem. However, we often want to use them for nonlocal approximations. The existing literature shows that Taylor series approximations are often quite

good even for states not close to the deterministic steady state. Judd and Guu (1993, 1997) investigated this issue in simple growth models. They find that, for empirically reasonable choices of tastes and technology, linear approximations do well for small but nontrivial neighborhoods of the deterministic steady state. However, the region of validity will be too small if the stochastic shocks cause the state to wander substantially from the deterministic steady state. Fortunately, they find that the quality of the approximations improve substantially as the higher-order terms are added. They also find that the certainty nonequivalence terms are important to achieve high quality approximations for stochastic approximations. More precisely, they substitute the computed Taylor series into the defining equations and evaluate the resulting error. The resulting error for capital stocks near the steady state is often the order of machine zero, an accomplishment which few other methods can claim. While their investigations have been limited to relatively small models, there is no reason to suspect that the performance of this approach will decay drastically as we move to larger models. In any case, any user of these methods should use some diagnostics to estimate the region where the constructed series is a good approximation.

It is tempting to compute higher-order approximations and then just assume that they are better than the certainty equivalent linear approximation. This approach is dangerous since it ignores the essential fact of Taylor series expansions – their range of validity is limited. Some elementary analysis shows the importance of this fact. Suppose that $f(x) = (9802 - 198x + x^2)^{-1}$ and we wanted to compute its power series around $x = 100$. The fourth-order Taylor series is

$$\widehat{f}(x) = 99990001 - 3999800x + 59999x^2 - 400x^3 + x^4$$

To measure the accuracy of this Taylor series, we computed the relative error in logarithm terms,

$$E(x) = \log_{10} \left| \frac{\widehat{f}(x)}{f(x)} - 1 \right|.$$

The results are displayed in Table 1.

Table 1: Relative errors in Taylor series expansion of $(9802 - 198x + x^2)^{-1}$						
$x - 100 :$.1	.2	.3	.5	1	1.5
$E(x) :$	-13	-9.7	-7.2	-4.2	0	2.4

Table 1 says that the fourth-order Taylor series approximation has very small errors for x within 0.5% of the central value of $x = 100$, but that it falls apart when $|x - 100| > 1$. We also computed the order 5, 10, and 20 Taylor series expansions and found the approximations to get better for $|x - 100| < 1$ but worse for $|x - 100| > 1$. It appears that we cannot construct a Taylor series based at $x = 100$ which is valid for any value of x more than 1% away from $x = 100$. The key fact is that the radius of convergence for the power series expansion of $f(x)$ around $x = 100$ is 1. This follows directly from the theory of power series in the complex plane. The polynomial $9802 - 198x + x^2$ has two zeroes of $x = 100 \pm \sqrt[2]{-1}$, both of which are distance 1 away from $x = 100$. Therefore, the infinite-order Taylor series based at $x = 100$ has a radius of convergence equal to 1. Radii of convergence for power series can be small; in fact, they can be arbitrarily small even when the function is simple and has no singularities on the real line. We have no idea about the radius of convergence for the Taylor series constructed by our methods. This is particularly problematic for us in stochastic models where, in reasonably calibrated cases, we do expect the state variables to roam more than 1% away from the deterministic steady state.

This cautionary example and the portfolio example above both show that we need ways to determine if our solutions are good, and these evaluations should be performed for any application before a computed approximation is accepted and used to make some economically substantive point. Therefore we need to develop diagnostic tools

which can be applied to any problem.

To measure the accuracy of our approximations we evaluate

$$E(x, \varepsilon) = \max_{\mu} \tilde{g}^{\mu} \left(x, \hat{Y}(x, \varepsilon), F(x, \hat{Y}(x, \varepsilon), \varepsilon z), \hat{Y}(F(x, \hat{Y}(x, \varepsilon), \varepsilon z), \varepsilon), \varepsilon z) \right)$$

where \tilde{g}^{μ} is a normalized, unit-free, version of g^{μ} . The specific details of the normalization depends on the equation. For example, Euler equations should be normalized so that the errors are in terms of percentage of consumption. Specifically, the unnormalized Euler equation is often expressed as

$$0 = u'(c_t) - \beta E \{ u'(c_{t+1}) R_{t+1} \}$$

which has units “utils per unit of consumption”. We need to get rid of both the utility and consumption units. A unit-free version is

$$0 = 1 - \beta \frac{E \{ u'(c_{t+1}) R_{t+1} \}}{u'(c_t)}$$

a form often used in empirical studies. If an equation is a market-clearing condition with form $0 = S - D$ for supply S and demand D , then a natural unit-free form would be $0 = 1 - D/S$ where any deviation expresses the excess supply as a fraction of total supply. In general, we need to use some set of equations \tilde{g}^{μ} where deviations from zero represent a unit-free measure of irrationality of the agents, lack of market-clearing, mistakes in predictions, and whatever else is involved in describing equilibrium, all of which we want to make small in any approximation.

7. SPECIFIC EXAMPLES

We have developed the full perturbation method for stochastic models and proposed diagnostic tests to ascertain accuracy. We next apply this approach to dynamic

programming problems of the form

$$\begin{aligned} \max_{\{L_t, I_{it}\}} \quad & E_t \sum_{\tau=t}^{\infty} \beta^{\tau-t} u(c_{\tau}, l_{\tau}) \\ c_t = & F(K_t^1, K_t^2, \dots, K_t^n, L_t, \theta_t) - \sum_{i=1}^n I_t^i \\ K_{t+1}^i = & K_t^i + \varphi_i \left(\frac{I_t^i}{K_t^i} \right) I_{it}, \quad i = 1, \dots, n \\ \theta_{t+1} = & \lambda \theta_t + \sigma \xi_{t+1} \end{aligned}$$

where K_t^i , $i = 1, \dots, n$, is the stock of type i capital goods at the beginning of period t , I_t^i is gross investment of type i capital in period t , θ_t is the productivity level in period t , and ξ_t is the i.i.d. productivity shock with mean zero and unit variance. We assume that ξ_t is truncated Normal with truncation at 3 standard deviations. The function $\varphi_i(I_t^i/K_t^i)$ represents net investment of type i capital after deducting adjustment costs. We assume that $\varphi_i(\cdot)$ has the following form

$$\varphi_i(s) = 1 - \frac{\delta_i}{2} s \quad (28)$$

Denote $K_t = (K_t^1, K_t^2, \dots, K_t^n)$, $s_t^i = I_t^i/K_t^i$. Let $V(K_t)$ be the value function. The Bellman equation for the above problem can be written as:

$$V(K_t) = \max_{L_t, s_t^i} u(c_t, l_t) + \beta E_t V(K_{t+1})$$

subject to

$$\begin{aligned} c_t &= F(K_t, l_t) - \sum_{i=1}^n s_t^i K_t^i \\ K_{t+1}^i &= (1 + \varphi_i(s_t^i) s_t^i) K_t^i, \quad i = 1, \dots, n \\ \theta_{t+1} &= \lambda \theta_t + \varepsilon \sigma \xi_{t+1} \end{aligned} \quad (29)$$

The law of motion equation in (29) displays the position and of the perturbation parameter ε . The ξ_t shocks are i.i.d. random variables and are unchanged by the perturbation. The deterministic case is when $\varepsilon = 0$ since then the ξ_t shocks have no effect on production. When $\varepsilon = 1$, the variance of the productivity shocks is σ^2 .

The first order conditions are:

$$0 = u_c(c_t, l_t)F_L(K_t, l_t) + u_l(c_t, l_t) \quad (30)$$

$$0 = -u_c(c_t, l_t) + \beta E_t \{V_i(K_{t+1})\} (\varphi'_i(s_t^i)s_t^i + \varphi(s_t^i)) \quad (31)$$

$$V_i(K_t) = u_c(c_t, l_t) (F_i(K_t, l_t) - s_t^i) + \beta E_t \{V_i(K_{t+1})\} (1 + \varphi_i(s_t^i)) \quad (32)$$

while $F_i = \partial F(K, L)/\partial K^i$, $V_i = \partial V(K)/\partial K^i$. Note that the Euler equations (30), (31) and (32) are functional equations, which implicitly defines the policy functions $l_t = L(K_t)$ and $s_t^i = S^i(K_t)$, and the gradient functions $V_i(K)$. We are going to solve these functions by the perturbation method as described by Judd (1998), that is, to compute Taylor expansions around the steady state and use them as approximations.

7.1. Error Bounds. At each capital stock K_t , the error bound of our solution, $E(K_t)$, is defined as the maximum of absolute Euler equation errors at this point.

$$E(K_t) = \max\{\|E_L\|, \|E_{K_i}\|, \|E_{V_i}\|, i = 1, \dots, n\}$$

where E_L , E_{K_i} , E_{V_i} are normalized errors, as given by:

$$E_L = (u_c F_L + u_l)/u_l$$

$$E_{K^i} = \beta E_t \{V_i(K_{t+1})\} (\varphi'_i(s_t^i)s_t^i + \varphi(s_t^i)) / u_c - 1$$

$$E_{V_i} = 1 - u_c(F_i - s_t^i) - \beta E_t \{V_i(K_{t+1})\} (1 + \varphi_i(s_t^i)) / V_i(K_t)$$

Normally we expect the error bounds become bigger as we move away from the steady state. To see how the errors grow, we introduce an overall measure of error bounds as a function of relative distance from the steady state. Formally, for every $r \geq 0$, we can define:

$$E(r) = \sup \left\{ E(K) \left| \sum_{i=1}^n \left(\frac{K^i - \bar{K}^i}{\bar{K}^i} \right)^2 \leq r \right. \right\}$$

where \bar{K}^i denotes the steady state value of K^i . The $E(r)$ is the error bound function we are seeking.

In practice, however, we cannot check all points on the surface of a sphere. We must confine to some finite sets. Let $D = \{-1, 0, 1\}^n$. Define

$$X = \left\{ (x^1, \dots, x^n) \left| x^i = \frac{d^i}{\sqrt{\sum_{j=1}^n (d^j)^2}}, d \in D, d \neq 0 \right. \right\}$$

Thus all points in X are on the surface of a unit sphere. We define our error bound function as

$$E(r) = \max\{ E(K) \mid K_i = (1 + rx^i)\bar{K}_i, x \in X \}$$

7.2. Computational Results. Our examples use the following functional forms:

$$\begin{aligned} u(c, l) &= \frac{c^{1-\gamma}}{1-\gamma} - \frac{l^{1+\eta}}{1+\eta}, \quad \gamma > 0, \eta > 0 \\ F(K, l) &= (K^1)^{\alpha_1} (K^2)^{\alpha_2} \dots (K^n)^{\alpha_n} l^{1-\sum_i \alpha_i}, \quad \alpha_i > 0, \sum_i \alpha_i < 1 \end{aligned}$$

We compute the model for the cases of 2,3 and 4 capital goods. In all the cases we choose $\beta = 0.95$, and $\delta_i = 0.1$, all i . We test the algorithm on several parameter values displayed in Table 1. For each combination of parameters in Table 1, we compute the first through fourth, and sometimes fifth, order Taylor series expansion. For each case, we compute $E(r)$ for various values of radii r , dimensions n , and expansion orders k . We then find the worst case for each scenario. That is, for radius r , dimension n , and order k , we find that case which had the worst $E(r)$. We report the worst cases in Table 2. For example, in the two capital good cases, the worst Euler equation error for the linear approximation at radius r was $10^{-3.2}$. That worst case may be different for the $r = .05$ case and for the $k = 2$ case. Therefore, every solution for the cases in Table 1 was better than the errors reported in Table 2.

Table 1: Parameter Values

(n, α_i) :	(2,.15), (3,.1), (4, .075)
γ :	0.5, 2, 5, 10
η :	10, 3, 1
(λ, σ) :	(0, 0), (0.05, 0), (0.10, 0) (0.01, 0.90), (0.01, 0.95)

The results in Table 2 show that the higher order approximations are very valuable and follow our intuition. For a fixed order k , the errors of the k 'th order approximation increase as we move away from the steady state. The linear approximation is acceptable for $r < .01$, but is of questionable value for $r > .05$. None of the approximations have acceptable Euler equation errors for $r = .5$.

For any fixed radius r we see that there is substantial payoff to using higher-order approximations. In particular, at $r = 0.10$, the linear approximation has Euler equation errors up to 10% of consumption but the fifth-order approximation has normalized errors on the order of 10^{-5} , an improvement of four orders of magnitude.

Table 2: Error bounds $\log_{10} E(r)$

r	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
2 capital good cases					
0.01	-3.2	-3.7	-4.1	-5.5	-5.7
0.05	-1.8	-2.8	-4.1	-5.1	-5.7
0.10	-1.1	-2.1	-3.1	-4.1	-5.1
0.20	-0.5	-1.2	-1.9	-2.6	-3.4
0.30	-0.1	-0.6	-1.3	-1.8	-2.3
0.40	0.2	-0.2	-0.7	-1.1	-1.6
0.50	0.6	0.2	-0.2	-0.6	-1.0
3 capital good cases					
0.01	-3.2	-3.8	-4.0	-5.5	
0.05	-1.8	-2.9	-4.0	-5.2	
0.10	-1.2	-2.2	-3.3	-4.3	
0.20	-0.6	-1.3	-2.1	-2.8	
0.30	-0.2	-0.7	-1.3	-1.9	
0.40	0.2	-0.3	-0.8	-1.3	
0.50	0.5	0.1	-0.4	-0.8	
4 capital good cases					
0.01	-3.3	-3.9	-4.1	-5.6	
0.05	-1.9	-3.0	-4.1	-5.6	
0.10	-1.3	-2.3	-3.4	-4.4	
0.20	-0.7	-1.4	-2.2	-2.9	
0.30	-0.3	-0.8	-1.5	-2.1	
0.40	0.1	-0.4	-0.9	-1.5	
0.50	0.4	-0.1	-0.5	-1.0	

Table 2 examined the quality of the approximations near the deterministic steady state. While this information is useful and exactly the kind of information which is related to the implicit function theorem, it does not tell us what we need to know about how good the approximation is for a stochastic problem because we do not know the range of the states. For example, the linear approximation looks good for only k within 5% of the steady state. If the capital stocks stay within that range, the linear approximation may be acceptable, but we would not be so accepting if the stochastic shocks pushes some capital stocks to levels more than 10% away from the steady state.

Tables 3 and 4 address these issues with stochastic simulation for a particular case. Tables 3 and 4 takes a degree k approximation and uses it to simulate the economy for 10^5 periods. We compute the deviation, $(K_t - \bar{K})/\bar{K}$, from the steady state and the magnitude of the Euler equation error at each realized state. Table 3 reports the mean deviation from the steady state of the capital stock, the standard deviation, and the maximum deviation. The mean deviation for $k = 1$ is nearly zero, as it should be since the linear approximation is a certainty equivalent approximation. Higher order approximations indicate that the mean capital stock is about 2% from the deterministic steady state, a fact not possible to approximate with the linear approximation. The other moments are largely unaffected by the higher orders of approximation.

Table 3: $(K - \bar{K})/\bar{K}$ along the simulation path

$$\gamma = 10, \eta = 10, (\sigma, \lambda) = (0.1, .95)$$

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
mean	-0.002	0.019	0.018	0.019	0.019
std. dev.	0.087	0.090	0.089	0.089	0.089
maximum	0.257	0.304	0.291	0.294	0.293
minimum	-0.249	-0.229	-0.227	-0.226	-0.226

Table 4 reports the mean of the absolute value of the errors, their standard deviation, and the maximum Euler equation error over the 10,000 period simulation. Since the true distribution of the states is not centered at the deterministic steady state, the results in Table 4 are not as impressive as in Table 2, but they again indicate the great value of higher-order approximations.

Table 4: Error bounds $\log_{10} E(r)$ along simulation paths

$$\gamma = 10, \eta = 10, (\sigma, \lambda) = (0.1, .95)$$

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$
mean	2.6×10^{-2}	1.2×10^{-3}	1.1×10^{-4}	6.2×10^{-5}	3.4×10^{-7}
std. dev.	3.7×10^{-2}	2.5×10^{-3}	1.5×10^{-3}	1.6×10^{-4}	1.4×10^{-6}
maximum	4.4×10^{-1}	4.0×10^{-2}	9.8×10^{-3}	3.0×10^{-3}	2.9×10^{-5}

We need to be clear about these error results. We do not present them to indicate that higher-order perturbation methods are good approximations and that the reader should feel free to apply them to his problems. Our point is that these error analyses need to be done as part of any application of perturbation methods. It is the critical fifth step in the perturbation method. The statistics displayed in Tables 2 and 4 should be reported in any application of the perturbation method just as t

statistics and confidence intervals are reported in any application of regression and other statistical methods.

Table 5 displays the computational costs associated with the higher-order approximations. We see that the number of derivatives to compute rise substantially as we increase approximation order and dimensions. There is a similar increase in time and space needed to compute the approximations. We include statistics on space since the space necessary to store all the necessary derivatives may be a limitation for perturbation methods. While the computational costs are substantial, they are not a serious problem. With increasing speed of computers and the fall in memory prices, perturbation methods are clearly competitive with alternatives for multidimensional problems.

Table 5: Computation costs

no. of capitals	state vars.	endog. vars.	order						
			1	2	3	4	5	6	7
number of derivatives to compute									
2^a	2	5	10	25	45	70	100	135	175
2^b	3	5	20	70	170	345	625	1045	
3^a	3	7	21	63	133	238	385		
3^b	4	7	35	140	385	875	1757		
4^a	4	9	36	126	306	621			
4^b	5	9	54	243	747	1881			
computing time in seconds									
2^a	2	5	0	0.03	0.37	3.38	23.6	148	923
2^b	3	5	0	0.06	0.83	15.5	199	907	
3^a	3	7	0	0.13	2.34	35.0	415	308	
3^b	4	7	0	0.22	6.58	127	1424		
4^a	4	9	0	0.40	9.89	202			
4^b	5	9	0	0.66	22.8	640			
memory used in megabytes									
2^a	2	5	2.5	2.5	2.8	4.2	12.0	48.0	200
2^b	3	5	2.6	2.6	3.1	7.2	51.0	440	
3^a	3	7	2.6	2.6	3.8	17.0	132		
3^b	4	7	2.7	2.7	4.8	33.0	386		
4^a	4	9	2.7	2.8	6.9	74.0			
4^b	5	9	2.8	2.9	9.5	135			

Note: a is the riskless case and b is the risky case

Table 6 summarizes the steps of the perturbation method.

Table 6: Perturbation Method for Rational Expectations Models

Step 1:	Compute the deterministic steady state with nonlinear equation solver
Step 2:	Compute linear approximation with some rational expectations solution method
Step 3:	Compute higher-order terms of deterministic problem through differentiation and linear equation solving
Step 4:	Compute stochastic deviation terms through differentiation and linear equation solving
Step 5:	Compute normalized errors in ergodic set of states through deterministic sampling and stochastic simulation

8. CONCLUSION

This paper has shown that it is feasible to apply perturbation methods to numerically solve rational expectations models substantially more complex than the usual representative agent, single good model. However, theory shows that the perturbation approach faces some limitations related to the range of the stochastic shocks and the local validity of the approximations. In response, we develop diagnostic methods to evaluate the approximations.

REFERENCES

- [1] Bensoussan, Alain. *Perturbation Methods in Optimal Control*, John Wiley and Sons, 1988.
- [2] Bishop, Richard L., and Samuel I. Goldberg, *Tensor Analysis on Manifolds*, Dover Publications, 1981.
- [3] Blanchard, Olivier Jean and Charles M. Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations" *Econometrica* **48**:1305–1311.
- [4] Brock, William A., and L. J. Mirman. 1972. "Optimal Economic Growth and Uncertainty: The Discounted Case," *Journal of Economic Theory* **4**: 479–513.
- [5] Campbell, John Y., and Luis M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press, 2002.
- [6] Chen, Baoline and Peter Zadrozny, "Higher Moments in Perturbation Solution of the Linear-Quadratic Exponential Gaussian Optimal Control Problem," Rutgers University, *Computational Economics*, forthcoming.
- [7] Christiano, Lawrence J. 1990. "Solving the Stochastic Growth Model by Linear-Quadratic Approximation and by Value-Function Iteration," *Journal of Business and Economic Statistics*, **8**:23–26.
- [8] Collard, Fabrice, and Michel Juillard, "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control* **25**, 2001, 979-999.
- [9] Collard, Fabrice, and Michel Juillard, "Perturbation Methods for Rational Expectations Models," Manuscript. Paris: CEPREMAP, February 2001.

- [10] Fleming, Wendell. 1971. "Stochastic Control for Small Noise Intensities," *SIAM Journal of Control* **9** (No. 3, August, 1971).
- [11] Fleming, W. H., and Souganidis, P. E. "Asymptotic series and the method of vanishing viscosity," *Indiana University Mathematics Journal* **35** (1986), no. 2, 425-447
- [12] Judd, Kenneth L. *Numerical Methods in Economics* , MIT Press, 1998.
- [13] Judd, Kenneth L. 1996. "Approximation, Perturbation, and Projection Methods in Economic Analysis," in *Handbook of Computational Economics*, edited by Hans Amman, David Kendrick, and John Rust, North Holland.
- [14] Judd, Kenneth L., and Sy-Ming Guu. 1993. "Perturbation Solution Methods for Economic Growth Models," in *Economic and Financial Modeling with Mathematics*, Hal Varian, ed., Springer-Verlag: New York.
- [15] Judd, Kenneth L., and Sy-Ming Guu. "Asymptotic Methods for Aggregate Growth Models," *Journal of Economic Dynamics and Control* **21** (No. 6, June 1997): 1025-42.
- [16] Kim, Jinill and Sunghyun Henry Kim, "Spurious Welfare Reversals in International Business Cycle Models," *Journal of International Economics*, forthcoming.
- [17] Kydland, Finn E., and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations," *Econometrica* **50**: 1345-1370.
- [18] Laitner, John. 1981. "The Stability of Steady States in Perfect Foresight Models," *Econometrica* **49**:319-333.
- [19] Magill, J. P. Michael. "A Local Analysis of N -Sector Capital Accumulation under Uncertainty," *Journal of Economic Theory*, **15** (1977): 211–219.

- [20] Marcet, Albert. “Simulation Analysis of Dynamic Stochastic Models: Applications to Theory and Estimation,” in C. Sims (ed.) *Advances in Econometrics Sixth World Congress*, Vol II., 91–118, Cambridge University Press.
- [21] McCullagh, P. (1987). *Tensor Methods in Statistics*. Chapman & Hall, London.
- [22] McGrattan, Ellen R. 1990. “Solving The Stochastic Growth Model by Linear-Quadratic Approximation.” *Journal of Business and Economic Statistics* **8**: 41–43.
- [23] Misner, Charles W., Kip S. Thorne, and John A. Wheeler, *Gravitation*, W.H. Freeman & Co, 1973.
- [24] Samuelson, Paul A. “The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances and Higher Moments,” *Review of Economic Studies* **37** (1970): 537–542.
- [25] Schmitt-Grohe, S, and M. Uribe. “Solving Dynamic General Equilibrium Models Using a Second-Order Approximation of the Policy Function, Working Paper, Rutgers University and University of Pennsylvania, 2002.
- [26] Sims, Christopher. “Second Order Accurate Solution Of Discrete-time Dynamic Equilibrium Models”, 2000, mimeo.
- [27] Taylor, John B., and Harald Uhlig. 1990. “Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods,” *Journal of Business and Economic Statistics*, **8**:1–18.
- [28] Tesar, Linda. “Evaluating the Gains from International Risksharing,” *Carnegie-Rochester Conference Series on Public Policy*, **42** (1995): 95-143. **5** (1995): 43-78.

- [29] Zeidler, E. *Nonlinear Functional Analysis and Its Applications: Volume I*, Springer-Verlag: New York, 1986.