Estimating the Accuracy of Numerical Solutions to Dynamic Optimization Problems

Michael Reiter

June 23, 2000

PRELIMINARY VERSION

Abstract

The paper considers stochastic dynamic optimization problems where a tentative numerical solution has been found. It uses the Euler residuals along simulated paths of the model to estimate the accuracy of the proposed solution. The main measure of accuracy is the the reduction in the criterion function from using the numerical rather than the exact solution. The method can also be used to estimate the approximation error of the policy function. The method is applied to the one-dimensional stochastic growth model, where it is shown to provide quite precise estimates of the errors in the value and the policy function.

Address of the author:

Michael Reiter
Universitat Pompeu Fabra
Department of Economics and Business
Ramon Trias Fargas, 25–27
08005 Barcelona
Spain

e-mail: reiter@upf.es

1 Introduction

With the necessity to solve higher dimensional optimization problems, economists resort more and more to approximative numerical methods whose accuracy properties are not well known. It therefore seems necessary to analyze the accuracy of a proposed solution in each specific case, to make sure that it is sufficient for the problem at hand. Often there are several numerical solution procedures available, and the task is to choose the best one.

For the evaluation of numerical solutions, it is natural to make use of the Euler residuals (Judd 1992). Since Euler equations are a fundamental tool in economic analysis, economists can probably interpret Euler residuals more easily than other error statistics. Den Haan and Marcet (1994) use the Euler residuals to construct a test for the null hypothesis that the numerical solution is identical to the exact solution. However, in that paper they do not try to measure the magnitude of the deviation between the exact and the numerical policy. Santos (1999) derives theoretical relationships between the size of the Euler residuals and the deviations of the policy function and the value function from their exact values. The derived relationship is in the *supremum norm* of these variables.

The present paper follows Santos (1999) in measuring the approximation error in the policy and the value function, but it is complementary to that paper in the sense that it does not estimate the supremum (or maximum) of the error over the state space, but either the error at a particular point in the state space or the average error over a part of the state space. In particular, the paper focuses on the loss in the objective function that results from using the numerical rather than the exact solution. While Santos (1999) mainly derives theoretical bounds that are easy to compute but relatively loose, the present paper uses methods that are more computer intensive but give tight estimates.

Computing average errors rather than error bounds in supremum norms is somewhat unusual and deserves some further discussion. Both the "supremum approach" and the "average approach" have their advantages. The supremum approach gives us the highest level of confidence in a method, if we can actually show that the supremum of the approximation error is below some required threshold level. This is probably the reason why it plays a dominant role in error analysis. The disadvantage is that the supremum approach may often be non-operational if the accuracy of the solution at hand is not very high. Accuracy measures in supremum norms are very conservative. First because the nature of supremum arguments makes that the estimated supremum error is often much bigger (by one or two orders of magnitude) than the true supremum error (cf. Section 2). Second, the supremum error criterion itself may be too demanding. Consider a numerical solution that provides the approximately right policy in almost all circumstances, but makes substantial mistakes under some circumstances. If these circumstances are very unlikely to ever occur, this numerical method may be easily good enough for a given purpose. In the same spirit, if we have to decide between two methods to use, it is not clear whether we should prefer the one with the lower supremum error or with the lower average error.

To clarify the latter point, it might be useful to distinguish the perspective of the agent who has to solve the problem (which I will call the "internal" perspective), from the perspective of a scientist who tries to describe, predict or interpret the behavior of an agent who solves such a problem. I will call this the "external" perspective. From

the internal point of view, the criterion of evaluation is clear: the method is better that obtains the higher value of the objective function, given the initial condition of the agent. If the same algorithm is used in many occasions, one should average the obtained value over the different (expected) initial conditions.

An external observer may have very different criteria to judge a solution. One such criterion may be the supremum norm. However, the application of these criteria requires that the observed agent itself (the object of study) uses an algorithm that is either exact or at least close to the exact algorithm *in terms of the external observer's criteria*. This is a strong requirement; if it is not met, or if the external observer does not have very clear criteria, it appears most natural to judge methods by the internal view.

The plan of the paper is as follows. Section 2 provides a heuristic discussion of the relationship between Euler residuals and approximation errors. Section 3 describes the class of optimization problems that we are going to study. Section 4 derives estimates of the approximation error in the value function, while the error in the policy function is treated in Section 5. Numerical examples are provided in Section 6, and Section 7 concludes.

2 Euler residuals and approximation errors

The method of the present paper makes essential use of the Euler residuals. The purpose of this section is therefore to gain some intuitive understanding about the relationship between the Euler residuals and the approximation error.

One potential problem is that the Euler equation alone is not a *sufficient* condition for a solution. An optimization problem usually has infinitely many paths that satisfy the Euler equation, i.e., have zero Euler residuals. Fortunately, this is not as big a problem as it seems, since the Euler equation together with a transversality condition are sufficient (Stokey and Lucas 1989, Theorem 4.15). This implies that a non-optimal solution which satisfies the Euler equation violates the transversality condition, and this means that, at some point in time, it will do something that is either very obviously non-optimal (in the consumption problem, letting the ratio of consumption to capital go to zero) or impossible (negative capital, growing at the rate of interest). Since a reasonable numerical method excludes both types of cases, the inaccuracy of the solution will *sooner or later* show up in the Euler residuals. We will detect the inaccuracy if we keep track of the Euler residuals over time.

Unfortunately, in order to get a precise estimate of the approximation error, it is not enough to look at the maximum Euler residual. Let us illustrate this at a very simple case. Consider the deterministic finite horizon consumption problem with interest rate and discount factor equal to zero. The household maximizes

$$\sum_{t=1}^{T} u(c_t) \tag{1}$$

with $u'(c) > 0$ and $u''(c) < 0$, subject to the constraint

$$\sum_{t=1}^{T} c_t = K \tag{2}$$

with given initial capital K. The Euler residual for this problem is

$$u'(c_t) - u'(c_{t+1}) \tag{3}$$

and the optimal consumption path is

$$c_t = c^* = K/T \tag{4}$$

Now assume that T is even and consider two small deviations from the optimal policy (cf. Figure 1)

$$C_t^1 = \begin{cases} c^* - \frac{1}{2}\epsilon & \text{for } t \text{ odd} \\ c^* + \frac{1}{2}\epsilon & \text{for } t \text{ even} \end{cases} \tag{5}$$

$$C_t^2 = c^* + \epsilon \left(t - \frac{T+1}{2} \right) \tag{6}$$

for small ϵ. Both policies satisfy the budget constraint (2). Up to a quadratic approxi-

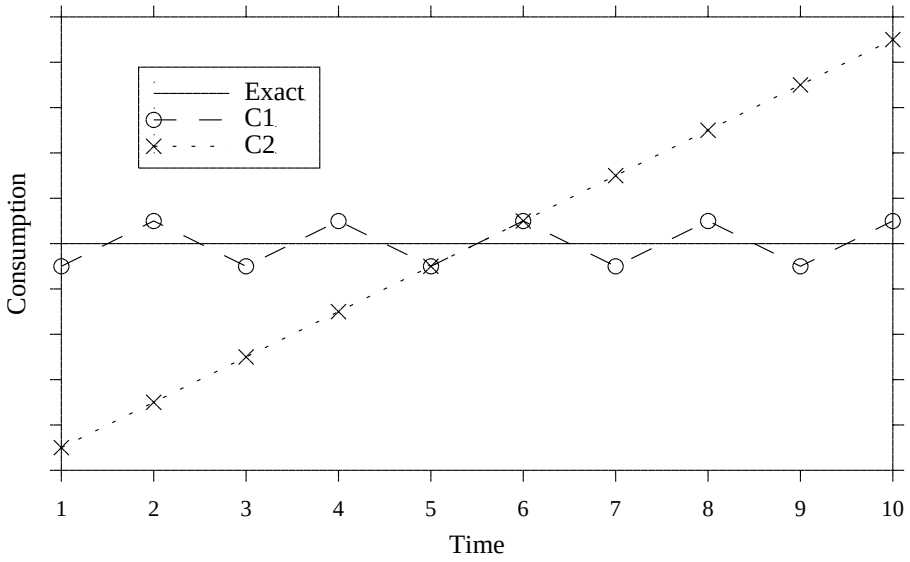

Figure 1: Example: finite horizon consumption problem

mation about c^*, the Euler residual for both policies and all t is given by $R_t = u''(c^*)\epsilon$, but the maximal approximation error of the policy function is 0.5ϵ for C_t^1 and $\frac{T-1}{2}\epsilon$ for C_t^2. The loss in value is $\frac{T}{8}|u''(c^*)|\epsilon^2$ for C_t^1 and $|u''(c^*)| \sum_{t=T/2+1}^{T} \epsilon^2 \left(t - \frac{T+1}{2}\right)^2$ for C_t^2.

The second policy has a much bigger error, because the mistake at a given point, as measured by the Euler residual, steadily adds up over time, while in the first policy the errors tend to cancel over time. This shows that there is no functional relationship between the maximum of the Euler residuals and the maximal approximation error. A theory that relates maximal Euler residuals to approximation errors will in most cases substantially overestimate the approximation error, since it has to account for the worst possible case (cf. the results in Santos, 1999, Section 4.3). If we want to obtain exact error estimates based on Euler residuals, we therefore have to monitor the Euler residuals

systematically over time. The next section will describe a method which does exactly this.

Note also that, with a given maximal Euler residual, the policy can deviate the more from the optimal policy, the longer the time horizon. In an infinite horizon problem, long time horizon should be interpreted as low discount factor, which explains the role of the discount factor in (Santos 1999, Theorem 3.3).

3 The Model

In this paper, we consider optimization problems of the form

$$\max_{u_0,u_1,\ldots} \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t F(x_t, u_t, z_t) \tag{7a}$$

subject to

$$x_{t+1} = u_t + g(z_{t+1}), \qquad \mathrm{E}_t\, g(z_{t+1}) = 0 \tag{7b}$$

$$u_t \in \Gamma(x_t, z_t) \tag{7c}$$

$$x_0 \in X \text{ given} \tag{7d}$$

and the transversality condition

$$\mathrm{E}_0 \lim_{t\to\infty} \beta^t F_1(x_t, u_t, z_t) \cdot x_t = 0 \tag{7e}$$

Here, x_t is a vector of endogenous state variables from the state space X, u_t is a vector of control variables, and z_{t+1} is a vector of exogenous random variables that satisfy a law of motion which need not be specified at the present stage. The random vector can appear in the objective function as well as in the dynamic equation, with the requirement that $\mathrm{E}_t\, g(z_{t+1}) = 0$. In other words, the dynamic equation (7b) is written such that the control u_t is the expected value of next period's state vector, which simplifies notation, The function F is assumed to be measurable, and three times continuously differentiable in (x_t, u_t) for all z_t. We also assume that F is strictly concave in x and u, which means that

$$H_t \equiv \begin{bmatrix} F_{11,t}(x_t, u_t, z_t) & F_{12,t}(x_t, u_t, z_t) \\ F_{21,t}(x_t, u_t, z_t) & F_{22,t}(x_t, u_t, z_t) \end{bmatrix} \tag{8}$$

is strictly negative definite for all (x_t, u_t). This implies that the Euler equation together with the transversality condition (7e) is sufficient for an optimum.

The set of feasible controls $\Gamma(x_t)$ is assumed to be convex for all x_t. In most of this paper we assume that the exact as well as the proposed numerical solution are inner solutions to this problem (cf. Santos, 1999, Section 2 for a discussion of this assumption). Section 4.4 will deal with the case of occasionally binding constraints.

We assume that the state space X is compact. Economic models are often defined on unbounded state spaces, but for numerical solutions it is usually necessary to rewrite the model in such a way that it can be well approximated on a bounded space. We take the bounded approximation as given and do not consider the approximation error that may result from approximating an unbounded problem on a bounded space. Since X is compact, the continuity of F and its derivatives implies that they are bounded. We also assume that the policy function is bounded.

Formulation (7) is of course not the most general stochastic optimization model, but it encompasses most of the models studied in macroeconomics. More general models do not always have an Euler equation, while for the above model we have the simple Euler equation

$$R_{t+1} \equiv F_2\left(x_t, u_t, z_t\right) + \beta \mathrm{E}_t\, F_1\left(x_{t+1}, u_{t+1}, z_{t+1}\right) = 0 \tag{9}$$

As argued above, the Euler residual R_{t+1} will not be exactly zero for numerical solutions. The numerical residuals will form the basis for the estimation of the errors in the value and policy function.

4 Estimating the error in the value function

We assume that a numerical solution for the problem (7a) is provided, and our task is to estimate the size of the approximation error of this solution. Of course, we do not assume to know the exact solution. We mostly concentrate on accuracy measured in terms of the objective function, more precisely, on the decrease in the objective function that results from using the approximate rather than the exact solution. Section 4.2 provides an upper bound to the error. In the numerical examples given below, we will se that this upper bound is rather tight. Section 5 will discuss the problem of measuring the error in the policy function.

4.1 Introduction: the deterministic case

This section discusses briefly and somewhat heuristically the deterministic case. This will provide some intuitive understanding of the accuracy estimates, even if we will later see that the analysis in the stochastic case differs substantially from the deterministic analysis.

Consider the deterministic version of the model with criterion function $F(x_t, u_t)$ and $x_{t+1} = u_t$. Assume we have simulated the path $\hat{x}_t$, for $t = 0, \ldots, T$, by the numerical solution, starting from $\hat{x}_0 = x_0$. Assume that the model is stable and that the simulated path converges at least approximately to the true stationary state. We can then choose T large enough so that what happens after time T is irrelevant, given a required degree of accuracy. The loss in accuracy from using the numerical rather than the exact solution is therefore approximately (means, to the required degree of accuracy) equal to the loss that we make in choosing a suboptimal path from x_0 to $\hat{x}_T$. This loss can be computed by solving the problem

$$\max_{\delta_1, \delta_2, \ldots} \sum_{t=0}^{T} \beta^t F(\hat{x}_t + \delta_{t-1}, \hat{x}_{t+1} + \delta_t) - F(\hat{x}_t, \hat{x}_{t+1}) \tag{10}$$

subject to

$$\delta_t + \hat{x}_{t+1} \in \Gamma(\delta_{t-1} + \hat{x}_t) \tag{11a}$$

$$\delta_{-1} = 0 \tag{11b}$$

$$\delta_T = 0 \tag{11c}$$

Here, δ_t is the time t policy error. The value function of this problem at x_0 gives us the value loss from the numerical solution.

Of course, problem (10) is just a finite horizon version of problem (7a) and not easier to solve. However, if the numerical solution $\hat{x}_t$ is already close to the true solution, the above problem can be substituted by the locally quadratic approximation

$$\max_{\delta_0,\delta_1,\ldots} \sum_{t=0}^{T} \beta^t \left[F_1'(\hat{x}_t, \hat{x}_{t+1})\delta_{t-1} + F_2'(\hat{x}_t, \hat{x}_{t+1})\delta_t + \frac{1}{2} \left({\delta_{t-1}}', {\delta_t}' \right) H_t \left(z^t \right) \begin{pmatrix} \delta_{t-1} \\ \delta_t \end{pmatrix} \right] \qquad (12)$$

where prime denotes transposition and $\hat{H}_t$ is the Hessian of $F(x_t, x_{t+1})$ at $(\hat{x}_t, \hat{x}_{t+1})$. This problem can be easily solved by backward induction, by the usual Riccati equations. The procedure gives an estimate of the policy error δ_{t-1} as well as the value loss from the approximation. (By iterating the quadratic approximation, we would probably converge to the exact solution, but this is not our concern here.)

From this we learn two things. First, computing the approximation error along a simulated path is a much simpler task than *solving* the optimization problem (imagine the problem is nonlinear and high-dimensional). Second, the accuracy with which we can estimate the approximation error depends itself on the accuracy of the numerical solution, because it determines the accuracy of the locally quadratic approximation (12). This will be made more precise in the next section.

4.2 General theory: an upper bound

From now on we deal with the general model of Section 3. Again, we follow the idea of estimating the approximation error along realized paths. One might first think that the analysis of the deterministic case can be carried over to the stochastic model by solving the stochastic version of the locally quadratic approximation along a realized path. This is not the case, for the following reason. The solution to a dynamic optimization model always has to take into account what happens off the realized path. In a deterministic model, it is sufficient to consider small deviations from the given path. In a stochastic model, the existence of (generally not small) random shocks force us to consider large deviations from a realized path, and these can, in general, not be well approximated by a quadratic function, even if we assume that the simulated path δ_t is close to the optimal path. The following computations will therefore be based on a more complicated argument; nevertheless, quadratic approximations will play a decisive role.

We first introduce some notation. A realized path of the numerical solution will always be denoted by $\hat{x}_t$. Such a path depends on the (given) initial value x_0 and a realization of the shocks $z_0, z_1, \ldots$. Let us denote the history of shocks $z_0, z_1, \ldots, z_t$ by z^t. Because of the Markov structure of the problem, the optimal policy u_t^* is a function of the vector of state variables. We also assume that the policy $\hat{u}_t$ provided by the numerical solution is a function of the state vector. This also implies that, with a fixed initial state x_0, the variables x_t^*, u_t^*, $\hat{x}_t$ and $\hat{u}_t$ are a function of the history of shocks z^t. All expectation operators should be understood as referring to the distribution of the shocks $z_0, z_1, \ldots$. Formally, the operator E_t means expectation w.r.t. the σ-algebra generated by z^t. We will say that a random variable is t-measurable to express that it is a function of z^t.

We now define the policy error $\delta_t \left(z^t \right) = u_t^* \left(z^t \right) - \hat{u}_t \left(z^t \right)$. The notation stresses the fact that we understand δ_t as a function of the history of shocks z^t, not of the current state vector. From (7b) it follows that the optimal state x_t^* satisfies $x_t^* = \hat{x}_t + \delta_{t-1}$. The first and second partial derivatives of the objective function along a simulated path

$\hat{x}_t$ are denoted by $F_{1,t}$, $F_{11,t}$ etc. They should also be understood as random variables depending on z^t.

Using this notation, we can now define the approximation error Δ of the value function as

$$\Delta \equiv \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t \Big[F\left(x_t^*, u_t^*, z_t\right) - F\left(\hat{x}_t, \hat{u}_t, z_t\right) \Big]$$

Analogous to the last section, consider the program

$$\Delta^{trunc} = \max_{\delta_1, \delta_2, \ldots} \sum_{t=0}^{T} \beta^t F(\hat{x}_t + \delta_{t-1}, \hat{x}_{t+1} + \delta_t, z_t) - F(\hat{x}_t, \hat{x}_{t+1}, z_{t+1}) \tag{13}$$

subject to (11a), (11b) and

$$\delta_T(z^T) = 0 \tag{14}$$

The difference between Δ and Δ^{trunc} satisfies

$$\|\Delta - \Delta^{trunc}\| \leq \mathrm{E}_0\, \beta^{T+1} \left[\|v\left(x_{T+1}^*\right) - v\left(\hat{x}_{T+1}\right)\| + \|v\left(\hat{x}_{T+1}\right) - w\left(\hat{x}_{T+1}\right)\| \right] \tag{15}$$

where $v(x)$ is the (exact) value function at point x, and $w(x)$ is the value reached under the policy of the numerical solution. The boundedness of F shows that this term goes to zero for $T \to \infty$, so we can concentrate on measuring Δ^{trunc} rather than Δ if we choose T large enough. We proceed in three steps. First, we derive a simple expression for Δ^{trunc}, based on a quadratic approximation. Second we show a way to compute an upper bound to this quadratic approximation based on an estimate of the policy error. Third, we show how to obtain an estimate of the policy error.

Step 1: Quadratic approximation to Δ^{trunc}

From now on, denote by δ_t the solution to program (13). Since the optimal policy is bounded, we can define

$$\eta = \sup_{t, z^t} \|\delta_t(z^t)\| \tag{16}$$

We now use a Taylor approximation of the terms in brackets about $\hat{x}_t$. Since the third partial derivatives of F are bounded, Taylor's theorem shows that

$$\Delta^{appr} \equiv \mathrm{E}_0 \sum_{t=0}^{T} \beta^t \Big[F_{1,t}\delta_{t-1} + F_{2,t}\delta_t + \frac{1}{2}\Big(\delta_{t-1}' F_{11,t} \delta_{t-1} + 2\delta_{t-1}' F_{12,t} \delta_t + \delta_t' F_{22,t} \delta_t \Big) \Big]$$
$$\geq \Delta^{trunc} - O\left(\eta^3\right) \tag{17}$$

In the following, we use the Euler equation to obtain a simpler expression for Δ^{appr}. Along any path z^{T+1}, the approximation errors δ_t satisfy

$$F_2\left(\hat{x}_t + \delta_{t-1}, \hat{u}_t + \delta_t, z_t\right) + \beta\, \mathrm{E}_t\, F_1\left(\hat{x}_{t+1} + \delta_t, \hat{u}_{t+1} + \delta_{t+1}, z_{t+1}\right) = 0 \tag{18}$$

Using again Taylor's theorem, we obtain the approximation

$$F_{2,t} + F_{21,t}\delta_{t-1} + F_{22,t}\delta_t + \beta\, \mathrm{E}_t \left(F_{1,t+1} + F_{11,t+1}\delta_t + F_{12,t+1}\delta_{t+1}\right) = O\left(\eta^2\right) \tag{19}$$

If we premultiply (19) by δ_t', use the fact that δ_t is a function of z^t so that $\delta_t' \mathrm{E}_t(x) = \mathrm{E}_t[\delta_t' x]$ (Billingsley 1986, Theorem 34.3) for any vector of random variables x, take expected values and use the law of iterative expectations, we get

$$\mathrm{E}_0 \Big[\delta_t' F_{2,t} + \delta_t' F_{21,t} \delta_{t-1} + \delta_t' F_{22,t} \delta_t + \beta \left(\delta_t' F_{1,t+1} + \delta_t' F_{11,t+1} \delta_t + \delta_t' F_{12,t+1} \delta_{t+1} \right) \Big] = O\left(\eta^3\right) \quad (20)$$

This expression can be used to simplify (17) if we apply the following reordering of terms:

$$\begin{aligned} &\sum_{t=0}^{T} \beta^t \left(\delta_{t-1}' F_{11,t} \delta_{t-1} + 2\delta_{t-1}' F_{12,t} \delta_t + \delta_t' F_{22,t} \delta_t \right) \\ &\quad = \sum_{t=0}^{T-1} \beta^t \left[\delta_t' F_{21,t} \delta_{t-1} + \delta_t' \left(F_{22,t} + \beta F_{11,t+1} \right) \delta_t + \beta \delta_t' F_{12,t+1} \delta_{t+1} \right] \\ &\qquad + \delta_{-1}' F_{11,0} \delta_{-1} + \delta_{-1}' F_{12,0} \delta_0 + \beta^T \left(\delta_T' F_{21,T} \delta_{T-1} + \delta_T' F_{22,T} \delta_T \right) \end{aligned} \quad (21)$$

If we multiply (20) by β^t, sum up for $t = 0, \ldots, T-1$, and subtract it from the summation term in (17), applying (21) and $\delta_{-1} = 0$, we see that

$$\begin{aligned} \Delta^{appr} = -\frac{1}{2} \mathrm{E}_0 \sum_{t=0}^{T} \beta^t \Big(&\delta_{t-1}' F_{11,t} \delta_{t-1} + \delta_{t-1}' F_{12,t} \delta_t + \delta_t' F_{21,t} \delta_{t-1} + \delta_t' F_{22,t} \delta_t \Big) \\ &+ \mathrm{E}_0 \beta^T \Big[F_{2,T} \delta_t + \Big(\delta_t' F_{21,T} \delta_{t-1} + \delta_t' F_{22,T} \delta_t \Big) \Big] + O\left(\eta^3\right) \end{aligned} \quad (22)$$

Therefore, if we choose β large enough so that the last terms in (22) are sufficiently small, our task reduces to estimating the quadratic term in (22), which we denote by Δ^{qu} and write as

$$\Delta^{qu} \equiv -\frac{1}{2} \mathrm{E}_0 \sum_{t=0}^{T} \beta^t \left({\delta_{t-1}}', {\delta_t}' \right) H_t \begin{pmatrix} \delta_{t-1} \\ \delta_t \end{pmatrix} \quad (23)$$

using the notation

$$H_t = \begin{pmatrix} F_{11,t} & F_{12,t} \\ F_{21,t} & F_{22,t} \end{pmatrix}$$

The concavity assumption made in Section 3 means that (23) is a positive definite quadratic form, and the estimated value gain Δ^{qu} is nonnegative.

Step 2: Computing an upper bound for Δ^{qu} using an estimate of δ_t

In the stochastic model, we cannot "compute" the δ_t's, not even for a given realization of z^T. This is apparent from Equ. (19), which can be slightly rearranged as

$$R_{t+1} + F_{21,t} \delta_{t-1} + \left(F_{22,t} + \beta \mathrm{E}_t F_{11,t+1} \right) \delta_t + \beta \mathrm{E}_t \left[F_{12,t+1} \delta_{t+1} \right] = O\left(\eta^2\right) \quad (24)$$

The Euler residual R_{t+1} can be measured, but the expectation $\beta \mathrm{E}_t \left[F_{12,t+1} \delta_{t+1} \right]$ cannot, because it requires the values of δ_{t+1} off the realized path. To compute δ_0 to δ_T recursively, the values of δ at all realizations starting from x_0 are required.

However, we will see below that we can use Equ. (24) to obtain an estimate of δ_t, which we denote by $\bar{\delta}_t$. If we compute formula (23) with $\bar{\delta}_t$ rather than δ_t, we obtain

$$\Delta^{ub} \equiv -\frac{1}{2}\,\mathrm{E}_0 \sum_{t=0}^{T} \beta^t \left(\bar{\delta}'_{t-1}, \bar{\delta}'_t\right) H_t \begin{pmatrix} \bar{\delta}_{t-1} \\ \bar{\delta}_t \end{pmatrix} \tag{25}$$

If we denote the error in estimating δ_t by $\tilde{\delta}_t = \delta_t - \bar{\delta}_t$, we can write

$$\Delta^{ub} = -\frac{1}{2}\,\mathrm{E}_0 \sum_{t=0}^{T} \beta^t \left(\left(\delta_{t-1} - \tilde{\delta}_{t-1}\right)', \left(\delta_t - \tilde{\delta}_t\right)'\right) H_t \begin{pmatrix} \delta_{t-1} - \tilde{\delta}_{t-1} \\ \delta_t - \tilde{\delta}_t \end{pmatrix} \tag{26}$$

$$= \Delta^{qu} + \Delta^{cross} - \Delta^{esterr} \tag{27}$$

where

$$\Delta^{cross} \equiv \mathrm{E}_0 \sum_{t=0}^{T} \beta^t \left(\delta_{t-1}{}', \delta_t{}'\right) H_t \begin{pmatrix} \tilde{\delta}_{t-1} \\ \tilde{\delta}_t \end{pmatrix} \tag{28}$$

$$\Delta^{esterr} \equiv \frac{1}{2}\,\mathrm{E}_0 \sum_{t=0}^{T} \beta^t \left(\tilde{\delta}'_{t-1}, \tilde{\delta}'_t\right) H_t \begin{pmatrix} \tilde{\delta}_{t-1} \\ \tilde{\delta}_t \end{pmatrix} \tag{29}$$

We therefore reach the conclusion that $\Delta^{ub} \geq \Delta^{qu}$ if we can show that Δ^{cross} is equal to zero. In other words, under this condition, Δ^{ub} is an upper bound for Δ^{qu}. Note that we can compute Δ^{ub} from (25) by Monte-Carlo techniques, to any desired degree of accuracy. The details of the computation are developed in Section 4.3.

Step 3: Estimating δ_t

To form an estimate $\bar{\delta}_t$ of δ_t, we start from (19). We replace $\mathrm{E}_t\, F_{1,t+1}$ by an estimate $F^e_{1,t+1}$ that has the property $\mathrm{E}_t\, F^e_{1,t+1} = \mathrm{E}_t\, F_{1,t+1}$. As discussed above,
$\mathrm{E}_t\,[F_{11,t+1}\delta_t + F_{12,t+1}\delta_{t+1}]$ is not available, so we simply ignore the expectation operator E_t and use $[F_{11,t+1}\delta_t + F_{12,t+1}\delta_{t+1}]$ (but note that the realization can be interpreted as an unbiased estimate of its expected value). Then we obtain the second order difference equation

$$F_{2,t} + F_{21,t}\bar{\delta}_{t-1} + F_{22,t}\bar{\delta}_t + \beta\left(F^e_{1,t+1} + F_{11,t+1}\bar{\delta}_t + F_{12,t+1}\bar{\delta}_{t+1}\right) = 0 \tag{30}$$

For a given realization of the series $\hat{x}_t$, we can compute the estimates $\bar{\delta}_t$ from Equ. (30) and the boundary conditions

$$\bar{\delta}_{-1} = 0 \tag{31a}$$

$$\bar{\delta}_T = 0 \tag{31b}$$

which are natural since δ_{-1} and δ_T are zero by construction. The following remarks should clarify some of the properties of this estimate.

1. The expression $F_{2,t} + \beta F^e_{1,t+1}$ is an estimate of the Euler residual $F_{2,t} + \beta\,\mathrm{E}_t\, F_{1,t+1}$. We do not require to measure the Euler residual precisely, but we will get a better estimate if we do so.

2. If the estimated Euler residual is zero for all t, the estimated error $\bar{\delta}_t$ is zero for all t.

3. We make an error in estimating δ_t mainly because we ignore the expectation operator and replace this part by the realized values. The error will be bigger, the larger is the variance of $[F_{11,t+1}\delta_t + F_{12,t+1}\delta_{t+1}]$ conditional on t.

4. $\bar{\delta}_t$ depends on the whole realization $x_0, \ldots, x_T$, so $\bar{\delta}_t$ is not t-measurable.

It remains to show that the error $\tilde{\delta}_t = \delta_t - \bar{\delta}_t$ has the property that the term Δ^{cross} in (28) is equal to zero. Subtracting (30) from (19) gives

$$F_{21,t}\tilde{\delta}_{t-1} + F_{22,t}\tilde{\delta}_t + \beta\,\mathrm{E}_t\left(F_{1,t+1} + F_{11,t+1}\delta_t + F_{12,t+1}\delta_{t+1}\right) - \beta\left(F^e_{1,t+1} + F_{11,t+1}\bar{\delta}_t + F_{12,t+1}\bar{\delta}_{t+1}\right) = 0 \quad (32)$$

Conditioning on t we see that

$$\mathrm{E}_t\left[F_{21,t}\tilde{\delta}_{t-1} + F_{22,t}\tilde{\delta}_t + \beta\left(F_{11,t+1}\tilde{\delta}_t + F_{12,t+1}\tilde{\delta}_{t+1}\right)\right] = 0 \quad (33)$$

Premultiplying (33) by δ_t', using again $\delta_t'\,\mathrm{E}_t(x) = \mathrm{E}_t\left[\delta_t' x\right]$, and taking expectations E_0, we get

$$\mathrm{E}_0\left[\delta_t' F_{21,t}\tilde{\delta}_{t-1} + \delta_t' F_{22,t}\tilde{\delta}_t + \beta\left(\delta_t' F_{11,t+1}\tilde{\delta}_t + \delta_t' F_{12,t+1}\tilde{\delta}_{t+1}\right)\right] = 0 \quad (34)$$

If we multiply (34) by β^t, sum up for $t = 0, \ldots, T-1$, and use the reordering (21), we obtain

$$\mathrm{E}_0 \sum_{t=0}^{T} \beta^t \left[\left(\delta_{t-1}', \delta_t'\right) H_t \begin{pmatrix} \tilde{\delta}_{t-1} \\ \tilde{\delta}_t \end{pmatrix}\right] = \beta^T \delta_T' F_{21,T}\tilde{\delta}_{T-1} \quad (35)$$

So Δ^{cross} is zero except for a term that goes to zero for large T (since δ_T and $\tilde{\delta}_{T-1}$ are bounded).

Finally, I would like to point out why we have to use $F_{11,t+1}$ and $F_{12,t+1}$ in formula (30) and not $E_t[F_{11,t+1}]$ and $E_t[F_{12,t+1}]$. Since neither $\bar{\delta}_t$ nor δ_{t+1} are t-measurable, both variables may be correlated with $F_{11,t+1}$ or $F_{12,t+1}$, and replacing these variables by their time t expectations would not allow us to derive formula (33).

The calculations of this section are summarized by the following theorem.

Theorem 1. *If $\bar{\delta}_t$ is defined by Equ. (30) and the boundary conditions (31), then*

$$\Delta^{ub} \geq \Delta - O\left(\eta^3\right) - K\beta^T$$

holds for some constant K.

Theorem 1 is the main theoretical result of this paper. It shows that even in the stochastic case, the estimated errors along realized paths can be used to estimate an upper bound of the approximation error of the value function. The numerical results of section Section 6.2 will show that these upper bounds are rather tight.

It may be useful to list all sources of error that come into play in estimating Δ by Δ^{ub}:

1. The error from the quadratic approximation of F about $\hat{x}_t$, and from the linear approximation of the Euler equation about $\hat{x}_t$, which together give an error which is $O\left(\eta^3\right)$.

2. The truncation error $K\beta^T$ from the finite horizon of the simulation.

3. The sampling error in the simulation.

4. Imprecise measurement of the Euler residual.

5. The quadratic term Δ^{esterr}.

The errors 3, 4 and 2 can be made arbitrarily small by increasing the computational effort, using longer time horizons, more simulations etc. The error 1 depends on the nonlinearity of the model. It is a cubic function of the maximal deviation of the simulation, so it disappears when the numerical solution goes to the exact solution. The error 5 cannot be decreased by doing more simulations etc., and it is quadratic in the maximal deviation, just as Δ^{ub} itself. It is therefore the most serious source of error.

4.3 Recursive computation of the error estimate

This section describes recursive formulas for the computation of Δ^{ub} for a given realization $x_0, \ldots, x_T$. The formulas are similar to the recursions of the linear-quadratic control problem. If we define

$$R^e_{t+1} = F_{2,t} + \beta F^e_{1,t+1} \tag{36}$$

we have from (30)

$$\bar{\delta}_t = -\left(F_{22,t} + \beta F_{11,t+1}\right)^{-1} \left(F_{21,t}\bar{\delta}_{t-1} + R^e_{t+1} + \beta F_{12,t+1}\bar{\delta}_{t+1}\right) \tag{37}$$

Since $\bar{\delta}_T = 0$, we get

$$\begin{aligned} \bar{\delta}_{T-1} &= -\left(F_{22,T-1} + \beta F_{11,T}\right)^{-1} \left(F_{21,T-1}\bar{\delta}_{T-2} + R^e_T\right) && (38)\\ &= a_{T-1} + A_{T-1}\bar{\delta}_{T-2} && (39) \end{aligned}$$

This can be used as the starting point of the following recursion. If

$$\bar{\delta}_{t+1} = a_{t+1} + A_{t+1}\bar{\delta}_t \tag{40}$$

is given with some vector a_{t+1} and some matrix A_{t+1}, it follows from (37) that

$$\bar{\delta}_t = a_t + A_t\bar{\delta}_{t-1} \tag{41}$$

with

$$\begin{aligned} a_t &= -\left(F_{22,t} + \beta F_{11,t+1} + \beta F_{12,t+1}A_{t+1}\right)^{-1} \left(R^e_{t+1} + \beta F_{12,t+1}a_{t+1}\right) && (42a)\\ A_t &= -\left(F_{22,t} + \beta F_{11,t+1} + \beta F_{12,t+1}A_{t+1}\right)^{-1} F_{21,t} && (42b) \end{aligned}$$

The recursions (42) define $\bar{\delta}_{t+s}$ as a function of $\bar{\delta}_t$ for $s > 0$. This allows us to define

$$\Delta^{ub}_t(\bar{\delta}_{t-1}) = -\frac{1}{2}\,\mathrm{E}_0 \sum_{s=t}^{T} \beta^{t-s} \left(\bar{\delta}_{s-1}\left(\bar{\delta}_{t-1}\right)', \bar{\delta}_s\left(\bar{\delta}_{t-1}\right)'\right) H_s \begin{pmatrix} \bar{\delta}_{s-1}\left(\bar{\delta}_{t-1}\right) \\ \bar{\delta}_s\left(\bar{\delta}_{t-1}\right) \end{pmatrix} \tag{43}$$

and compute it recursively, starting with

$$-2\Delta^{ub}_T(\bar{\delta}_{T-1}) = \bar{\delta}'_{T-1}F_{11,T}\bar{\delta}_{T-1} \tag{44}$$

If we are given

$$-2\Delta_{t+1}^{ub}(\bar{\delta}_t) = \bar{v}_{t+1} + v_{t+1}'\bar{\delta}_t + \bar{\delta}_t' V_{t+1}\bar{\delta}_t \tag{45}$$

it follows that

$$\begin{aligned}-2\Delta_t^{ub}(\bar{\delta}_{t-1}) &= \bar{\delta}_{t-1}' F_{11,t}\bar{\delta}_{t-1} + 2\bar{\delta}_{t-1}' F_{12,t}\left(a_t + A_t\bar{\delta}_{t-1}\right) + \left(a_t + A_t\bar{\delta}_{t-1}\right)' F_{22,t}\left(a_t + A_t\bar{\delta}_{t-1}\right) \\ &\quad + \beta\left[\bar{v}_{t+1} + v_{t+1}'\left(a_t + A_t\bar{\delta}_{t-1}\right) + \left(a_t + A_t\bar{\delta}_{t-1}\right)' V_{t+1}\left(a_t + A_t\bar{\delta}_{t-1}\right)\right] \\ &= \bar{v}_t + v_t'\bar{\delta}_{t-1} + \bar{\delta}_{t-1}' V_t\bar{\delta}_{t-1}\end{aligned} \tag{46}$$

with

$$\bar{v}_t = a_t' F_{22,t} a_t + \beta\left[\bar{v}_{t+1} + v_{t+1}' a_t + a_t' V_{t+1} a_t\right] \tag{47}$$

$$v_t = 2F_{12,t}a_t + 2A_t' F_{22,t} a_t + \beta\left[A_t' v_{t+1} + 2A_t' V_{t+1} a_t\right] \tag{48}$$

$$V_t = F_{11,t} + 2F_{12,t}A_t + A_t' F_{22,t} A_t + \beta A_t' V_{t+1} A_t \tag{49}$$

Since $\bar{\delta}_{-1} = 0$, we finally arrive at the error estimate

$$\Delta_0^{ub} = -\frac{1}{2}\bar{v}_0 \tag{50}$$

The estimate Δ^{ub} in (26) is then the sum of Δ_0^{ub} over many simulated paths. To compute the error at a specific point of the state space, the simulations will all start from this point. The compute the (weighted) average of the error over parts of the subspace, we start the simulations from different points, according to some probability distribution over the state space.

4.4 Occasionally binding constraints

The calculations of Sections 4.2 and 4.3 can be carried over to the case of binding constraints in the following way. Assume that in the numerical solution, a constraint for the control variable i at time t is binding. We accept this and set $\bar{\delta}_t^i = 0$. For the components that do not face a binding constraint, the Euler equation is still valid. This leads to simple changes in the recursive formulas (42), but the calculation of Δ^{ub} is otherwise unchanged.

Since the assumption $\bar{\delta}_t^i = 0$ is not necessarily true, we introduce a new form of error here. Each instance of such an error is of the order $O\left(\eta^2\right)$. However, if η goes to zero, the probability of making this error also goes to zero, linearly in η, and so the claim that the error in Δ^{ub} is of order $O\left(\eta^3\right)$ still holds up.

[DETAILS TO BE COMPLETED.]

5 Estimating the error in the policy function

The error in the policy function is much more difficult to measure than the error in the value function. In the procedure of the last section we actually computed an estimate $\bar{\delta}_t$ of the policy error, but there is no theorem showing that this estimate is unbiased or an upper bound of the true policy error. In estimating the policy error we make a mistake, which we denoted by $\tilde{\delta}_t$, and everything we know about it is Equ. (33), which does not allow to draw inferences about the expectation or variance of $\tilde{\delta}_t$. The main obstacle to

further analysis is that we cannot say anything about the correlation between $\tilde{\delta}_{t+1}$ and $F_{12,t+1}$ (remember that the policy error is an arbitrary function of the state variables).

In the limiting case where the Hessian matrix is constant, Equ. (33) gives a linear second order difference equation in $\mathrm{E}(\tilde{\delta}_t)$, and with the boundary condition it implies $\mathrm{E}(\tilde{\delta}_t) = 0$ for all t. This also suggests that for smooth problems, where the Hessian moves slowly over time, the estimate is almost unbiased.

In the general case, one can follow two strategies to obtain an estimate of the policy error:

1. If we only want to know the approximate size of the policy errors and do not need rigorous error bounds, we can take the mean value of the estimated policy error (perhaps plus 2 standard deviations, to be on the safe side) as an estimate. The numerical experiments with variants of the stochastic growth model (cf. Section 6.1) suggest that it is a good estimate.

2. If one is willing to make a higher computational effort than that of Section 4.2 (but still less than the one needed to obtain the solution of the optimization problem), it is possible to compute the policy error at a certain point up to $O\left(\eta^2\right)$. This will be shown in the rest of this section.

To compute the estimate of the policy error, we analyze Equ. (24), with the $O\left(\eta^3\right)$-term set to zero, which gives

$$R_{t+1} + F_{21,t}\delta_{t-1} + \left(F_{22,t} + \beta\,\mathrm{E}_t\,F_{11,t+1}\right)\delta_t + \beta\,\mathrm{E}_t\left[F_{12,t+1}\delta_{t+1}\right] = 0 \tag{51}$$

Solving this for δ_t we obtain

$$\delta_t = -\left(F_{22,t} + \beta E_t[F_{11,t+1}]\right)^{-1}\left(F_{21,t}\delta_{t-1} + R_{t+1} + \beta\,\mathrm{E}_t\left[F_{12,t+1}\delta_{t+1}\right]\right) \tag{52}$$

Adopting again the approximation $\delta_T = 0$, we get

$$\delta_{T-1} = a^*_{T-1} + A^*_{T-1}\delta_{T-2} \tag{53}$$

where

$$a^*_{T-1} = -\left(F_{22,T-1} + \beta E_{T-1}[F_{11,T}]\right)^{-1} R_T \tag{54}$$

$$A^*_{T-1} = -\left(F_{22,T-1} + \beta E_{T-1}[F_{11,T}]\right)^{-1} F_{21,T-1} \tag{55}$$

This can be used as the starting point of the following recursion. Assume we are given the relationship

$$\delta_{t+1} = a^*_{t+1} + A^*_{t+1}\delta_t \tag{56}$$

where the vector a^*_{t+1} and the matrix A^*_{t+1} are known. Then it follows from (52) that

$$\delta_t = a^*_t + A^*_t\delta_{t-1} \tag{57}$$

where a^*_t and A^*_t are given by

$$a^*_t = -\left(F_{22,t} + \beta\,\mathrm{E}_t\left[F_{11,t+1} + F_{12,t+1}A^*_{t+1}\right]\right)^{-1}\left(R_{t+1} + \beta\,\mathrm{E}_t\left[F_{12,t+1}a^*_{t+1}\right]\right) \tag{58a}$$

$$A^*_t = -\left(F_{22,t} + \beta\,\mathrm{E}_t\left[F_{11,t+1} + F_{12,t+1}A^*_{t+1}\right]\right)^{-1} F_{21,t} \tag{58b}$$

Following this recursion we finally arrive at $\delta_0 = a_0^*$. Iterating Equs. (58) forwards, we obtain the policy error as an expectation of future Euler residuals:

$$\delta_0 = a_0^* = \mathrm{E}_0 \sum_{s=0}^{T-1} C_s R_{s+1} \tag{59}$$

where

$$C_0 \equiv F_{12,0}^{-1} M_0 \tag{60}$$

$$C_{t+1} \equiv \beta C_t M_{t+1} \tag{61}$$

$$M_t \equiv -F_{12,t} \left(F_{22,t} + \beta \, \mathrm{E}_t \left[F_{11,t+1} + F_{12,t+1} A_{t+1}^* \right] \right)^{-1} \tag{62}$$

From (58b) we see that M_t satisfies the recursion

$$M_t = -F_{12,t} \left(F_{22,t} + \beta \, \mathrm{E}_t \left[F_{11,t+1} + M_{t+1} F_{21,t+1} \right] \right)^{-1} \tag{63}$$

If we can compute the matrices M_t, we can use (59) and Monte-Carlo techniques to compute the policy error δ_0 (or at least an unbiased estimate). The precision of the estimate can be arbitrarily increased by increasing the number of Monte-Carlo simulations, subject to an $O\left(\eta^3\right)$-error from (51).

Comparing the recursions (58) with (42), we see that the difference lies in taking expectations as of time t. This is the reason why the estimate of δ_0 obtained here is unbiased, unlike its counterpart $\bar{\delta}_t$ of Section 4.2. This comes at a computational cost, however: while the recursions (42) can be computed along simulated paths, Equs. (58) are functional equations that have to be computed for the whole state space. Since the recursion (63) converges backwards in time, we can actually replace the M_t's in (59) by the solution to the time-invariant equation

$$M(x) = -F_{12,x} \left(F_{22,x} + \beta \, \mathrm{E}_x \left[F_{11,x'} + M(x') F_{21,x'} \right] \right)^{-1} \tag{64}$$

where we write all the expressions as functions of the state vector x, and x' denotes next periods state vector. Equ. (64) is similar to a Bellman equation, or other functional equations used in solving dynamic optimization problems. It is considerably simpler than those equations since it does not involve any optimization. It can be obtained either recursively on a finite grid, or by projection methods.

6 Numerical examples

This section reports numerical results on a well known test problem, the one-dimensional stochastic neoclassical growth model. Numerical solutions to variants of this model have been intensively studied in the literature. I have chosen a one-dimensional model, since this can be solved at a finite grid with so high precision that it can serve as the "exact solution" for test purposes (Santos and Vigo-Aguiar 1998, Section 4).

6.1 One-dimensional stochastic growth

The social planner maximizes

$$\mathrm{E} \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma} - 1}{1-\gamma} \tag{65}$$

subject to the constraints

$$\begin{aligned} c_t &= \theta A k_t^\alpha - k_{t+1} \\ k_t &\geq 0 \end{aligned}$$

where θ_t is an i.i.d. shock with log-uniform distribution

$$-\sigma \leq \log \theta_t \leq \sigma, \qquad \sigma > 0 \tag{66}$$

We bring this model into the form of Section 3 and solve it on a one-dimensional grid by defining

$$x_t = \log(\theta A k_t^\alpha) \tag{67}$$

$$u_t = \log(A k_{t+1}^\alpha) \tag{68}$$

$$z_t = \log \theta_t \tag{69}$$

$$g(z) = z \tag{70}$$

$$F(x_t, u_t, z_t) = \frac{\left[e^{x_t} - (e^{u_t}/A)^{1/\alpha}\right]^{1-\gamma} - 1}{1-\gamma} \tag{71}$$

The simulations use the parameter values $\beta = 0.95$, $\alpha = 0.4$, and different values of σ. A wide range of values for γ has been used, but always with $\gamma > 1$, which is necessary to make the function F in concave in x and u. The constant A was set to $1/(\beta\alpha)$ so that the deterministic steady state capital stock is equal to 1.

The theory developed above is based on quadratic approximations, and the results are derived up to an error term $O\left(\eta^3\right)$. In practice, it might then not be clear whether the estimates are sufficiently accurate if there are strong nonlinearities with a sharply changing second derivatives of the relevant functions. Since the stochastic growth model is known to be very well behaved, one might question the relevance of the results for less well-behaved problems.

To deal with this problem without investigating more complicated models, I make the growth model as nonlinear as possible, first by using very high values of the risk aversion parameter (up to $\gamma = 10$), and then also by introducing a kink in the utility function at a value c^* close to the steady state level of consumption. More precisely, for $c \leq c^*$ the risk aversion parameter is some γ_0, for $c > c^*$ it is some γ_1, and at $c = c^*$ the utility function is scaled such that the utility function is continuous and differentiable. The second derivative, however, suddenly jumps downward at this point. This violates the assumptions made in Section 3, but can be considered as the limit case of a model with rapidly changing second derivative. The switch point c^* was set at 1.02 times the deterministic steady state level of consumption.

6.2 Results

The solution was developed on a grid that conforms roughly to capital values $k \in (0.2, 5)$. The grid size for the "exact solution" was 3137, for the approximate solutions grid sizes of 5 (!) and 25 were used. Euler residuals of the growth model were calculated at a grid of 41 points.

The results are based on 20000 Simulations of 200 periods each. The shocks were not generated randomly, but as a sort of "subrandom" sequences, namely generalized Faure points (Papageorgiou and Traub 1996). This should increase the precision of

the estimates. The large number of simulations is mainly necessary to compute the difference between the value obtained with the exact and with the approximate solution. The estimate of the error bound cold have been obtained with fewer simulations, cf. Section 6.3.

Some results are presented in the Table 1. The results refer not to average errors,

γ	σ	Gridsize	Δ^{ub}	Δ	Est.Pol.Err.	True Pol.Err
2	0.2	5	3.649e-04	3.823e-04	-2.009e-03 (2.611e-03)	-2.101e-03
2	0.2	25	2.409e-06	3.142e-06	-1.077e-04 (2.758e-04)	-2.024e-04
2	1e-5	5	2.322e-05	2.313e-05	1.715e-03 (1.715e-03)	1.707e-03
10	0.2	5	5.039e-03	6.473e-03	-2.758e-02 (3.701e-02)	-3.789e-02
2,10	0.2	5	9.904e-03	6.750e-03	1.273e-02 (2.146e-02)	1.099e-02
2,10	0.2	25	4.231e-05	2.442e-05	1.996e-04 (6.647e-04)	2.084e-04

Notes: Two values of γ indicate kinked utility function
Gridsize: refers to approximate solution
Δ is exact loss in value function, Δ^{up} is estimated upper bound
Est.Pol.Err: mean (in parentheses: absolute mean plus 2 stdev) of estimated $\bar{\delta}_0$

Table 1: Simulation results, growth model

but to the error starting from a specific value of the capital stock, namely with capital 20 percent above the steady state. Two conclusions emerge from the table. First, the upper bound is relatively tight, it overestimates the error by at most 100 percent. Sometimes the upper bound is slightly below what is called the "true error", but one should keep in mind that the "true error" is also estimated and contains a sampling error component.

Second, the mean of the estimated policy error is close to the true error, so the estimate of $\bar{\delta}_0$ is almost unbiased in these examples. However, since the estimated error can be positive or negative, or close to zero by accident, it seems safer to use the statistic "absolute mean plus 2 standard deviations" of estimated error as a practical upper bound to the error in the policy function.

The results reported here always started from the same starting point. Numerous other runs using different starting values confirm the above conclusions.

[PRELIMINARY. MORE EXAMPLES WILL FOLLOW.]

6.3 Computational cost

The computational cost of calculating the upper bound to the value function error is not trivial, but it is substantially lower than the cost of calculating the numerical solution itself, at least for medium or higher-dimensional problems. This is because we only have to solve a series of linear-quadratic problems along realized paths, and calculate Euler residuals. The main effort lies in fact in the computation of the Euler residuals, and so the computational burden grows with the dimension of the state space in the same way as the burden of multi-dimensional integration. Choosing the right grids for integration (quadrature grids) will therefore prove essential for efficient computation. One should recall that the theory did not depend on calculating the Euler residuals correctly: less precisely calculated residuals will simply result in a not so tight error bound, cf. Section 4.2. The effort to produce these estimates therefore depends on the required precision on the upper bound. This is true also because it determines the

number of simulations used in computing the expectation in Δ^{ub}. In my examples, the upper bound could be computed up to an error of about 5 percent with about 200 simulations.

7 Conclusions

The paper has derived estimates of the approximation error in the value as well as the policy function of a numerical solution to a dynamic optimization problem, whereby the exact solution is not known. Test applications have shown that the error can be estimated with good precision.

References

Billingsley, P.: 1986, *Probability and Measure*, 2nd edn, John Wiley & Sons, New York.

Den Haan, W. J. and Marcet, A.: 1994, Accuracy in simulations, *Review of Economic Studies* **61**(1), 3–17.

Judd, K. L.: 1992, Projection methods for solving aggregate growth models, *Journal of Economic Theory* **58**(2), 410–52.

Papageorgiou, A. and Traub, J.: 1996, Beating Monte Carlo, *Risk* **9**(6), 63–5.

Santos, M. S.: 1999, Accuracy of numerical solutions using the euler equation residuals, University of Minnesota.

Santos, M. S. and Vigo-Aguiar, J.: 1998, Analysis of a numerical dynamic programming algorithm applied to economic models, *Econometrica* **66**(2), 409–26.

Stokey, N. L. and Lucas, R. E.: 1989, *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, MA.