

# Nonlinear and stable perturbation-based approximations

Wouter J. Den Haan and Joris De Wind\*

December 17, 2010

## Abstract

Users of regular higher-order perturbation approximations can face two problems: policy functions with odd undesirable shapes and simulated data that explode. Kim, Kim, Schaumburg, and Sims (2008) propose an alternative, namely pruned perturbation, which avoids the instability problem. In this paper, we document that pruned perturbation approximations have some important drawbacks. We propose an alternative perturbation-based approximation that (i) does not have odd shapes, (ii) generates stable time paths, and (iii) avoids the drawbacks that hamper pruning. We consider models for which the highlighted problems of regular higher-order perturbation are relevant. We find that our alternative and pruned perturbation approximations give a good qualitative insight in the nonlinear aspects of the true solution, but—with a few exceptions—differ from the true solution in some quantitative aspects, especially during severe peaks and troughs.

*Key Words:* accuracy, nonlinear numerical solutions

*JEL Classification:* C63, E21

---

\*Den Haan: University of Amsterdam and CEPR, e-mail: wdenhaan@uva.nl. De Wind: De Nederlandse Bank and University of Amsterdam, e-mail j.de.wind@dnb.nl. We would like to thank Fabio Canova, Ken Judd, Michel Juillard, Ondra Kamenik, Giovanni Lombardo, Tarik Ocaktan, the editor Paul Klein, and two anonymous referees for useful comments.

# 1 Introduction

Perturbation has become a popular choice to solve dynamic stochastic general equilibrium (DSGE) models. Unfortunately, regular higher-order perturbation approximations are not guaranteed to generate non-explosive time paths. Moreover, regular perturbation approximations are polynomials and the unavoidable oscillations of polynomials imply that higher-order approximations do not inherit properties such as monotonicity and convexity from the true underlying policy functions. This is a problem facing all approximation procedures that use polynomials as basis functions. The problem is especially severe for perturbation approximations, because perturbation analysis does not give the user the tools to relocate these problems to areas of the state space that are of no importance.<sup>1</sup> Consequently, the undesirable oscillations could occur close to the steady state.

To understand the problem, consider the following policy function:

$$x = f(x_{-1}) = \alpha_0 + x_{-1} + \alpha_1 e^{-\alpha_2 x_{-1}}, \quad (1)$$

with  $\alpha_0 = 0.3311$ ,  $\alpha_1 = 0.9$ , and  $\alpha_2 = 1$ . The true policy function has a unique fixed point (at  $x = 1$ ) and the dynamics are globally stable. Figure 1 plots this policy function and the second-order perturbation approximation.<sup>2</sup>

At the fixed point, the second-order perturbation approximation inherits three key properties of the true policy function: (i) increasing in  $x_{-1}$ , (ii) strictly convex in  $x_{-1}$ , and (iii) the approximation is locally stable, that is,  $(\partial f(x_{-1})/\partial x_{-1})|_{x=1} < 1$ . For *any* second-order polynomial with these properties, it must be true that the function value goes to  $+\infty$  as  $x_{-1}$  goes to  $+\infty$ . This means that the second-order perturbation approximation must have a second intersection with the 45° line, which in turn implies that the dynamics of the approximation are not globally stable.

---

<sup>1</sup>In contrast, the user of projection methods does have this type of control by choosing the appropriate grid. Typically, undesirable oscillations occur outside the grid, which means that one can push these oscillations out of the relevant area by widening the grid. Moreover, when using Chebyshev nodes one even obtains uniform convergence within the grid. See Chapter 6.5 in Judd (1998) for a discussion.

<sup>2</sup>When perturbation analysis is applied to DSGE models, then the derivatives of the unknown policy function are only implicitly defined. In this example, we know the (derivatives of the) policy function and the perturbation approximation is simply the Taylor-series expansion of  $f(x_{-1})$ .

For the policy function defined in Equation 1, the location of the second intersection of the perturbation approximation with the  $45^\circ$  line moves towards the steady state as  $\alpha_2$  increases. If the second intersection is sufficiently far away from the steady state, then the instability will have no practical consequences. In a non-stochastic environment, instability would then only occur when the initial values for  $x_{-1}$  is far away from the steady state. In a stochastic environment, the problematic part of the state space would only be reached in the case of extremely unlikely events. In this paper, we will document, however, that these types of problems cannot be ignored in practice.

Kim, Kim, Schaumburg, and Sims (2008) and Lombardo (2010) propose to use pruned perturbation to deal with the problem of exploding simulated data. Pruning is already used in several academic papers.<sup>3</sup> The standard  $n^{\text{th}}$ -order perturbation approximation uses *one*  $n^{\text{th}}$ -order policy function to generate *one* time path. In contrast, the pruning procedure generates *multiple* time paths using the time paths generated by lower-order approximations as the input for the higher-order terms in higher-order approximations. This modification does not alleviate the problem that higher-order perturbation approximations can have undesirable odd shapes. In fact, pruned perturbation approximations have several additional unattractive features. In particular, the pruned perturbation approximation is not a function of the original set of state variables. The pruning procedure introduces additional state variables and is thus a function of a larger set of variables. We argue that this makes several standard exercises at least a lot more cumbersome. Another striking feature of the pruned perturbation procedure is that the  $n^{\text{th}}$ -order perturbation approximation does not deliver an exact fit if the truth is an  $n^{\text{th}}$ -order polynomial even though pruned perturbation approximations are polynomials. This questions the suitability of pruned perturbation approximations when the underlying function is close to a low-order polynomial.

We propose an alternative procedure, perturbation-plus, that generates stable time paths, does not generate policy functions with odd shapes, and avoids the problems of

---

<sup>3</sup>See, e.g., Andreasen (2008), Fahr and Smets (2008), Doh (2009), and Fernández-Villaverde, Guerrón-Quintana, Rubio-Ramírez, and Uribe (2009).

pruning. It starts out with a first-order perturbation approximation. To solve for the period- $t$  model outcomes, we use the *exact* equations of the model for period  $t$  and for the  $J - 1$  subsequent periods. The system is closed by the assumption that the behavior in period  $t + J$  is determined by the first-order perturbation approximation. Although the procedure is easy to program, it is computing intensive unless the chosen value for  $J$  is low.

To evaluate whether the pruning and the perturbation-plus approximations are accurate, we consider (i) models where the parameter values are just below or just above the critical levels for which regular second-order perturbation approximations generate stable time paths and (ii) models for which regular second-order perturbation generates time paths that reach that part of the state space where the second-order approximation (of a monotone increasing function) has reached its problematic decreasing part. These are non-trivial numerical problems, which is not surprising given that regular second-order perturbation approximations face difficulties.<sup>4</sup> Neither the pruning nor the perturbation-plus approximations are in general very accurate, although we found some cases where the perturbation-plus approximation is accurate. Although the two modifications to the standard perturbation procedure do not always pass the accuracy tests with flying colors, the news is not all bad. Both procedures provide a good *qualitative* insight in how the true time paths differ from the path generated by first-order perturbation. That is, the methods provide a reasonable idea about the nonlinear effects, which are substantial in the models considered. Of course, there is no guarantee that these results carry over to other models and the user should be careful in using these procedures, especially when the nonlinearities are so important that regular higher-order perturbation approximations generate exploding series.

The organization of this paper is as follows. In Section 2, we describe the models we use to highlight the problems of higher-order perturbation and the drawbacks of the pruning procedure. In Section 3, we explain the problems of higher-order perturbation.

---

<sup>4</sup>Although these are non-trivial models, it is straightforward to obtain accurate approximations using projection methods since the number of state variables are small.

In Section 4, we discuss the pruning procedure and its drawbacks as well as our alternative, the perturbation-plus approximation. In Section 5, we evaluate the accuracy of the two perturbation-based approximations that always generate stable time paths. The last section concludes.

## 2 Models

In this section, we describe the models used to illustrate the problems of higher-order perturbations and the properties of the proposed approximations. The first model is the neo-classical growth model. The second model is also a very simple model, namely a representative-agent business cycle model in which the labor market is modelled using the Pissarides matching framework. We consider parameterizations of the model such that the volatility of aggregate employment relative to the volatility of aggregate labor productivity takes on plausible values. It is well known that this requires the match surplus to be sufficiently volatile.<sup>5</sup> The higher volatility of the surplus makes it more likely that the nonlinear features of the model matter. We use the matching model to illustrate that higher-order perturbation approximations are hampered with undesirable features *even* if the underlying model is a very simple representative-agent model in which the standard deviations of key aggregates like employment and output take on plausible, i.e. low, values.

The third model we consider is a (simple) model with heterogeneous agents. Agents face idiosyncratic income shocks and are constrained in their ability to insure themselves against these shocks, because they have no access to financial assets with payoffs contingent on the realization of the idiosyncratic shock and because they face transaction costs when trading the existing financial asset.

---

<sup>5</sup>See Hagedorn and Manovskii (2008).

## 2.1 Neo-classical growth model

The representative agent maximizes

$$\max_{\{c_t, k_t\}_{t=1}^{\infty}} \mathbb{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \frac{c_t^{1-\gamma} - 1}{1-\gamma}$$

s.t.

$$c_t + k_t = e^{z_t} k_{t-1}^{\alpha} + (1 - \delta)k_{t-1}, \quad (2)$$

$$z_t = \rho_z z_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_z^2), \quad (3)$$

$k_0, z_1$  given.

The Euler equation is given by

$$1 = \mathbb{E}_t \left[ \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} \left( \alpha e^{z_{t+1}} k_t^{\alpha-1} + (1 - \delta)k_t \right) \right]. \quad (4)$$

If  $\gamma = \delta = 1$ , then the model reduces to the Brock-Mirman model. This model is unusual in the sense that there are analytical solutions for the two policy functions. They are given by

$$k_t = \alpha \beta e^{z_t} k_{t-1}^{\alpha} \text{ and} \quad (5)$$

$$c_t = (1 - \alpha \beta) e^{z_t} k_{t-1}^{\alpha}. \quad (6)$$

## 2.2 Matching model

There are two types of agents in the model, workers and entrepreneurs. Both types of agents are members of a representative household. At the end of the period, the household receives wages and firm profits from its members. These are distributed equally among the household's members for consumption.

**Firms.** In this model, the key decision is made by a representative entrepreneur. The entrepreneur maximizes the discounted value of future firm profits. That is,

$$\max_{\{v_t, n_t\}_{t=1}^{\infty}} E_1 \sum_{t=1}^{\infty} \beta^{t-1} \left( \frac{c_t}{c_1} \right)^{-\gamma} ((e^{z_t} - w) n_{t-1} - \psi v_t)$$

s.t.

$$n_t = (1 - \rho_n) n_{t-1} + p_{f,t} v_t, \quad (7)$$

$$z_{t+1} = \begin{cases} z_t & \text{with probability } \tilde{\rho}_z \\ -z_t & \text{with probability } (1 - \tilde{\rho}_z) \end{cases} \quad (8)$$

$$n_0, z_1 \text{ given.} \quad (9)$$

Here,  $v_t$  is the amount of vacancies posted by the firm,  $\psi$  the cost of posting a vacancy,  $p_{f,t}$  is the number of matches per vacancy,  $p_{f,t} v_t$  is the total number of new hires, and  $c_t$  is the consumption level of the representative household. The wage rate,  $w$ , is assumed to be fixed.<sup>6</sup> The firm takes the value of  $p_{f,t}$  as given. The value of  $z_t$  can take on two values, namely  $-\zeta$  and  $+\zeta$ .

The first-order conditions are given by

$$\psi = p_{f,t} \lambda_t \text{ and} \quad (10)$$

$$\lambda_t = \beta E_t \left[ \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (e^{z_{t+1}} - w + (1 - \rho_n) \lambda_{t+1}) \right], \quad (11)$$

where  $\lambda_t$  is the Lagrange multiplier of the constraint that describes the law of motion of  $n_t$ . It represents the value generated when adding one extra worker to the firm's workforce.

**Consumers.** The representative household simply consumes the income earned by its members. Thus,

$$c_t = w n_{t-1} + (e^{z_t} n_{t-1} - w n_{t-1} - \psi v_t) = e^{z_t} n_{t-1} - \psi v_t. \quad (12)$$

**Matching market.** The number of new hires is determined on a matching market at which the  $1 - n_{t-1}$  workers that are not employed and firms search for a match. The total

---

<sup>6</sup>The assumption of sticky wages is helpful in generating sufficiently volatile employment.

number of matches,  $m_t$ , is given by

$$m_t = \phi_0(1 - n_{t-1})^\phi v_t^{1-\phi}, \quad (13)$$

which means that the number of matches per vacancy are given by

$$p_{f,t} = \phi_0 \left( \frac{1 - n_{t-1}}{v_t} \right)^\phi. \quad (14)$$

**Equilibrium.** Equations (7), (10), (11), (12), and (14) form a system of five equations per period that determine  $n_t$ ,  $v_t$ ,  $\lambda_t$ ,  $c_t$ , and  $p_{f,t}$  as a function of  $n_{t-1}$  and  $z_t$ .

**Effect of a shock.** An increase in the expected values of future productivity levels leads to an increase in  $\lambda_t$  and raises the expected benefit of posting a vacancy,  $p_{f,t} \lambda_t$ , above its cost,  $\psi$ . This leads to an increase in vacancies and employment until  $\psi$  and  $p_{f,t} \lambda_t$  are equal again.

**Keeping the problem smooth.** The variable  $p_{f,t}$  is typically interpreted as a matching probability and is restricted to be less than or equal to 1. This implies that the policy function is no longer smooth. The reason is the following. When  $z_t$  takes on very low values, then  $\lambda_t < \psi$ . If  $\lambda_t < \psi$ , then the value of an extra employee is less than the posting cost. If  $p_{f,t}$  is restricted to be less than 1, then it is impossible to satisfy Equation (10). In itself this is not a problem. It simply means that firms post no vacancies, that is, the firm is at a corner solution. Perturbation analysis can no longer be used, however, if firms occasionally hit corners.

To avoid this dilemma, we do *not* interpret  $p_{f,t}$  as a probability and we allow  $p_{f,t}$  to exceed 1. If  $p_{f,t}$  exceeds 1, then  $m_t > v_t$  and firms simply hire more than one worker on each posted vacancy. Equations (10) and (14) make clear that there always is an internal solution for  $v_t$  as long as

$$\lambda_t > 0, \quad (15)$$

a condition that is satisfied in our calibrated model.

Nondifferentiabilities also occur if  $m_t > u_t$ , but this possibility turns out to be irrelevant in our numerical analysis.



## 2.3 Modified Deaton model

The last model considered is a simple partial equilibrium model in which agents face idiosyncratic income risk. The agent solves the following optimization problem:

$$\begin{aligned} \max_{\{c_t, a_t\}_{t=1}^{\infty}} \quad & \mathbb{E}_1 \sum_{t=1}^{\infty} \beta^{t-1} \left( \frac{c_t^{1-\gamma} - 1}{1-\gamma} - P(a_t) \right) \\ \text{s.t.} \quad & \\ & c_t + \frac{a_t}{1+r} = a_{t-1} + e^{z_t}, \\ & z_t = \bar{z} + \varepsilon_t \text{ and } \varepsilon_t \sim N(0, \sigma_z^2), \\ & a_0 \text{ given.} \end{aligned} \tag{16}$$

Here,  $c_t$  stands for the agent's consumption level,  $a_t$  stands for the amount of assets chosen in period  $t$ ,  $z_t$  is an exogenous random income component, and  $r$  is the exogenous interest rate. There are two reasons why markets are not complete. First, there is only one financial asset, namely a risk-free bond. Second, there are transactions costs associated with trading in the financial asset, which we model as utility costs.

The model is very similar to the model in Deaton (1991), except that Deaton (1991) does not have transactions costs, but a non-negativity constraint on  $a_t$ . That is, Deaton (1991) assumes that

$$a_t \geq 0. \tag{19}$$

We specify our transactions cost or penalty function such that this inequality constraint is a special case of the model. In particular, the penalty function,  $P(a_t)$  is given by

$$P(a_t) = \frac{\eta_1}{\eta_0} \exp(-\eta_0 a_t) + \eta_2 a_t. \tag{20}$$

The value of  $\eta_0$  controls the curvature of the penalty function and the nonlinearity of the problem. Suppose that  $\eta_2 = 0$ .<sup>7</sup> Then

$$\lim_{\eta_0 \rightarrow \infty} p(a_t) = \begin{cases} \infty & \text{for } a_t < 0 \\ 0 & \text{for } a_t \geq 0 \end{cases}. \tag{21}$$

---

<sup>7</sup>The term  $\eta_2 a_t$  gives additional flexibility, which we exploit in the calibration. The additional term also makes it possible to ensure that the penalty term is equal to zero in the steady state, which may be convenient in some applications.

That is, as  $\eta_0 \rightarrow \infty$  our penalty function implements the standard non-negativity constraint on  $a_t$ .

Perturbation techniques cannot deal with inequality constraints. The penalty function, thus, does not only make the model more general, it also makes it possible to use perturbation techniques.<sup>8</sup>

The Euler equation is given by

$$\frac{c_t^{-\gamma}}{1+r} + \frac{\partial P(a_t)}{\partial a_t} = \beta \mathbb{E}_t \left[ c_{t+1}^{-\gamma} \right]. \quad (22)$$

### 3 Higher-order perturbation in practice

In the introduction, we mentioned two potential problems of higher-order perturbation, namely undesirable shapes and instability. The question arises whether these problems matter, that is, whether they occur in the relevant part of the state space in practice. Of course, the problems do not occur if the shocks hitting the system are sufficiently small.<sup>9</sup> But in practice one solves models to shed light on real world problems, which means one has to use a realistic amount of volatility.

In this section, we shed light on this question by analyzing the perturbation approximations of the solutions to the models of Section 2. Before discussing the results, we describe the characteristics of a perturbation approximation.

---

<sup>8</sup>De Wind (2008) compares the properties of the model with the inequality constraint given in Equation (19) with the corresponding properties of the model with the penalty function for different values of  $\eta_0$ . Not surprisingly, the value of  $\eta_0$  matters a lot for the tails of the generated distribution of  $a_t$ . In particular, the tail of the distribution generated with the model with a penalty function only matches the tail of the model with the non-negativity constraint for high values of  $\eta_0$ . But model properties such as the volatility of consumption depend a lot less on the value of  $\eta_0$ .

<sup>9</sup>In fact, one can use linear approximations if the shocks are sufficiently small.

### 3.1 Characteristics of perturbation approximations

DSGE models can typically be written as follows:

$$0 = E_t [H(x_t, x_{t-1}, y_{t+1}, y_t, z_{t+1}, z_t)], \quad (23)$$

$$z_{t+1} = \Lambda z_t + \varepsilon_{t+1}, \quad (24)$$

$$\varepsilon_{t+1} \sim N(0, \sigma\Omega). \quad (25)$$

Here,  $x_{t-1}$  is an  $n_x \times 1$  vector containing the state variables of the system,  $y_t$  is an  $n_y \times 1$  vector containing other endogenous variables that are not state variables,  $z_t$  is an  $n_z \times 1$  vector with the exogenous random variables,  $\varepsilon_t$  is the vector with the corresponding innovations,  $\sigma$  is a scalar that controls the overall volatility of the model ( $\sigma \geq 0$ ), and  $\sigma\Omega$  is the  $n_z \times n_z$  covariance matrix of the innovations. Finally,  $H(\cdot)$  is a vector-valued function with dimension  $n_x + n_y$ .

We denote the true rational expectations solution by

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} f(x_{t-1}, z_t; \sigma) \\ g(x_{t-1}, z_t; \sigma) \end{bmatrix} \quad (26)$$

and the  $n^{\text{th}}$ -order perturbation approximation by

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \tilde{f}_{n^{\text{th}}}(x_{t-1}, z_t; \sigma) \\ \tilde{g}_{n^{\text{th}}}(x_{t-1}, z_t; \sigma) \end{bmatrix}. \quad (27)$$

Let  $s_t$  denote the arguments of the policy function, that is,  $s_t = [x_{t-1}, z_t; \sigma]$  and let  $\bar{s}$  denote the corresponding steady state values, that is,  $\bar{s} = [\bar{x}, 0; 0]$ . The numerical approximation  $\tilde{f}(\cdot)_{n^{\text{th}}}$  is an  $n^{\text{th}}$ -order perturbation approximation if the following conditions hold:

$$f(x_{t-1}, z_t; \sigma)|_{s_t=\bar{s}} = \tilde{f}_{n^{\text{th}}}(x_{t-1}, z_t; \sigma)|_{s_t=\bar{s}} \quad (28)$$

$$\left. \frac{\partial^i f_{n^{\text{th}}}(x_{t-1}, z_t; \sigma)}{\partial x_{t-1}^j \partial \sigma^{i-j}} \right|_{s_t=\bar{s}} = \left. \frac{\partial^i \tilde{f}_{n^{\text{th}}}(x_{t-1}, z_t; \sigma)}{\partial x_{t-1}^j \partial \sigma^{i-j}} \right|_{s_t=\bar{s}} \quad \text{for } \begin{matrix} i = 1, \dots, n \\ j = 0, \dots, i \end{matrix}$$

A similar set of conditions determines whether  $\tilde{g}_{n^{\text{th}}}(\cdot)$  is an  $n^{\text{th}}$ -order approximation.

The true policy functions and the numerical approximations are functions of the endogenous state variables,  $x_{t-1}$ , and the exogenous state variables,  $z_t$ . This property is an important aspect of recursive models. Our formulations of both the true solution and the regular perturbation approximation allow for the possibility that a particular state variable has no effect on a particular choice, but it does not allow for variables other than the state variables to have an effect. Although this is a standard property of rational expectations solutions and numerical approximations, we highlight this property because, as shown below, the pruning approximation does not satisfy this property.

### 3.2 Perturbation approximations and the neo-classical growth model

It is well known that both log-linear and linear approximations are accurate for the neo-classical growth model as long as  $\sigma_z$  takes on plausible values or values that do not exceed those plausible values by too much.<sup>10</sup> In fact, the solution of the Brock-Mirman model is a log-linear function of capital and productivity. Consequently, *any* perturbation approximation will recover the true rational expectations solution if the model is written in the logarithms of the variables. To ensure that the problem remains non-trivial, even when we consider the Brock-Mirman version, we calculate approximations in the levels of the state variables, not in the logarithms.

Consistent with the facts stated in the last paragraph, we find that the higher-order terms of higher-order perturbation approximations are quantitatively small for parameter values commonly used in the literature. The perturbation approximations are then not affected by the problems discussed in the introduction. Therefore, we also consider values for  $\sigma_z$  and  $\gamma$  that are higher than those normally used. Although the objective of this section is to analyze the problems of higher-order perturbation approximations researchers encounter in practice, it is still useful to consider these not so typical parameter values. There are two reasons. First, the simplicity of the model makes it easy to understand why higher-order perturbation approximations run into problems. A better understanding of

---

<sup>10</sup>In terms of the other parameters, it is important that the coefficient of relative risk aversion,  $\gamma$ , is not too high.

the problems in such a simple case is helpful in understanding the problems in more complex cases. Second, the analysis points out that the highlighted problems eventually show up for some parameter values, even in very simple and almost linear models.

**Non-monotonicity of second-order perturbation approximations.** Panel A of Figure 2 plots the perturbation approximation of the capital policy function when  $\sigma_z$  is equal to 0.007, a very standard value. The policy functions are plotted as a function of  $k_{t-1}$  for three different values of  $z_t$ .<sup>11</sup> The results are shown for the Brock-Mirman model, that is,  $\gamma = \delta = 1$ . The other parameter values take on standard values.<sup>12</sup>

It is impossible for standard second-order perturbation approximations to be monotonically increasing (or decreasing). As documented by the figure, the problematic decreasing part occurs, however, when the capital stock is very high, namely when it is more than 2.5 times its steady state value. Simulated values for the capital stock would not reach such high values when  $\sigma_z = 0.007$ . Obviously, such high values could be reached when  $\sigma_z$  takes on higher values. The policy functions will change, however, if  $\sigma_z$  changes. It could be the case that the undesirable decreasing part of the policy functions is pushed to the right as  $\sigma_z$  increases. This would make it more difficult to reach the decreasing part. This turns out *not* to be the case, as is documented in Panel B of Figure 2. This panel plots the policy functions when  $\sigma_z$  is equal to 0.2. For the high value of  $z$ , the turning point after which the function is decreasing in  $k_{-1}$  is indeed pushed to the right. For the low value of  $z$ , however, the turning point is pushed to the left. In fact, the policy function at the low productivity value is downward sloping for a large range of values of  $k_{-1}$ .

**Instability of second-order perturbation approximations.** According to the second-order perturbation approximation, the choice for capital and consumption are strictly positive at  $k_{-1} = 0$ . In fact, the chosen levels are quite high. Of course, this is not accurate at all.<sup>13</sup> Although inaccurate, this property does make the perturbation approximation more

<sup>11</sup>The values are 0 and plus and minus two times the standard deviation of  $z$ .

<sup>12</sup>In particular,  $\alpha = 0.36$ ,  $\beta = 0.99$ , and  $\rho_z = 0.95$ .

<sup>13</sup>According to the true rational expectations solution, both capital and consumption should be set equal to 0 when the beginning-of-period capital stock equals 0; this is the only feasible choice.

stable than the true solution. When using the second-order perturbation approximation, one could even start at somewhat *negative* capital levels and the economy will still revert back to the area around the steady state.<sup>14</sup> We find this to be true for a wide range of parameter values, also when we move beyond the Brock-Mirman model and consider other values for  $\gamma$  and allow for partial depreciation.<sup>15</sup> Nevertheless, the instability problem of the second-order perturbation approach is just around the corner. This is discussed next.

The second-order perturbation approximation grossly violates the budget constraint for low values of the beginning-of-period capital stock. But it is easy to obtain a second-order perturbation approximation that exactly satisfies the budget constraint. In particular, one could use the perturbation approximation for either capital or consumption and use the budget constraint to solve for the other variable. The numerical approximation satisfies the conditions in Equation (28) and is, thus, a perturbation approximation.

Figure 3 plots the perturbation approximation for capital,  $\tilde{f}_{2nd}(k_{-1}, 0)$ , when consumption is determined by the standard second-order perturbation approximation and capital is solved from the budget constraint. The value of  $z$  is set equal to 0. The policy function then gives the dynamics of the system if there are no shocks.<sup>16</sup> The solid line corresponds to the case when the parameters are the same as those used to create Panel A of Figure 2.

Interestingly, the perturbation approximation is now monotonically increasing for all positive values of  $k_{-1}$ . Thus, in terms of avoiding odd shapes, this alternative is an improvement. In terms of stability it is not. The figure documents that the policy function

---

<sup>14</sup>The second-order perturbation solution for  $k_t$  does have a second fixed point, but the value of  $k$  at this second fixed point is negative. If the capital stock would ever get below this negative value, then the second-order perturbation solution would be explosive. Note that  $\lim_{k_{-1} \rightarrow \infty} \tilde{f}_{2nd}(k_{-1}, z) = -\infty$ . That is, if the economy would start out at (or reach) a sufficiently high capital stock, then the capital choice could be so negative that the solution gets into the unstable region. Given that the capital choice only turns negative at very high values of the existing capital stock, this is not something one would have to worry about.

<sup>15</sup>We could find a case with a second positive-valued fixed point, but we had to raise the value of  $\gamma$  to 35 and the value of  $k$  at this second fixed point is very small. In this case the economy would diverge for positive initial values when these initial values are below this second intersection.

<sup>16</sup>This policy function is also indicative of the *expected* dynamics if there are shocks.

for capital has an additional *positive*-valued fixed point to the left of the steady state.<sup>17</sup> The time paths that start out or reach such low capital levels are diverging.

If  $\gamma$  is equal to 1, then this second fixed point of the perturbation approximation occurs at a value of the capital stock that is only 10% of the steady state value. The figure also plots the second-order perturbation approximation for the case when  $\gamma$  is equal to 10. In this case, the value of capital at the second fixed point is substantially higher, namely around 30% of the steady state.

**Summary for the neo-classical growth model.** In addition to the cases discussed here, we have considered the properties of second-order perturbation approximations for several parameter values including those that allow for partial depreciation. We found that odd shapes and instability only happen in unlikely cases. That is, in practice second-order perturbation approximations of the neo-classical growth model do not exhibit the problems we highlighted in the introduction. Nevertheless, we do not consider the analysis here very comforting. It is true that one has to go outside the usual range of parameter values to encounter problems. But we have shown that higher-order perturbation approximations do have undesirable shapes and instability problems, even though the model is very simple and has a log-linear solution. This brings up the question whether the problems will occur for more standard parameter values when more interesting models are considered. We document in the next two subsections that we have to answer this question in the affirmative.

### 3.3 Perturbation approximations and the matching model

In this subsection, we consider the matching model. We will show that the standard second-order perturbation approximation generates explosive time paths when the model is calibrated to generate a realistic amount of volatility. As discussed in Section 2.2, the productivity level is assumed to be a discrete-valued random variable that can take on

---

<sup>17</sup>In fact, there is a third positive fixed point. But this fixed point is quite far away from the steady state and is ignored in the text. But if the economy would start out at capital levels above this third fixed point, then capital is expected to grow without bound.

only two values. The main advantage of this assumption is that it allows us to portray the reasons behind the problems of higher-order perturbation approximations with a simple graphical analysis. The other advantage of this assumption is that it is easy to ensure that the true model solution is always well-defined. In particular, we assume that the low value of  $z_t$  is such that profits are always positive.<sup>18</sup>

**Parameter values for the matching model.** We choose  $\zeta$  and  $\tilde{\rho}_z$  such that the standard deviation and the autocorrelation of  $z_t$  are equal to the corresponding numbers for the data generating process for  $z_t$  used to analyze the neo-classical growth model. This means that  $\zeta = 0.0224$  and  $\tilde{\rho}_z = 0.975$ . We set  $\beta = 0.99$  and  $\phi = 0.5$ .<sup>19</sup> The values of  $\psi$ ,  $\rho_x$ , and  $\phi_0$  are chosen such that the steady state values of the unemployment rate, the number of matches per unemployed worker, and the number of matches per vacancy, are equal to 5%, 0.7, and 0.7, respectively.<sup>20</sup> The value of  $\gamma$  is set equal to 4.5. This value for  $\gamma$  is perhaps a bit higher than the most commonly used values, but still a plausible value.<sup>21</sup> We consider two values for the wage rate, namely  $w = 0.96$  and  $w = 0.973$ . When  $w$  is equal to 0.96, then the volatility of the employment relative to the volatility of productivity is equal to 0.25, whereas the observed ratio in the data is equal to 0.437.<sup>22</sup> Thus, to match the observed relative volatility, the value of  $w$  has to be increased. When  $w = 0.973$ , then the volatility of employment relative to labor productivity (calculated

---

<sup>18</sup>This condition is sufficient to keep the problem well-defined, but it is not necessary. That is, we could allow profits to be somewhat negative. To keep the problem well-defined it is important that  $\lambda_t > 0$ . As long as expected future profits offset current losses, then  $\lambda_t$  would be positive.

<sup>19</sup>See Petrongolo and Pissarides (2001) for a motivation for the chosen value of  $\phi$ .

<sup>20</sup>This implies that  $\psi = 0.5965$ ,  $\rho_n = 0.368$ , and  $\phi_0 = 0.7$ .

<sup>21</sup>For lower values of  $\gamma$ , the second-order perturbation solution is still well-behaved when the volatility of employment relative to the volatility of productivity is equal 0.437, i.e., the observed value. For somewhat higher values of the target, the second-order perturbation approximation runs into the same problems as those discussed here. For example, when  $\gamma = 3.5$  and  $w = 0.976$ , then the relative volatility of employment would be equal to 0.50 according to the accurate projection solution and the 2<sup>nd</sup>-order perturbation approximation is no longer stable.

<sup>22</sup>The series are filtered using the HP filter. See Den Haan and Kaltenbrunner (2009) for further details on the data used.



using an accurate projection method) matches its observed counterpart.<sup>23</sup>

**second-order perturbation when  $w = 0.96$ .** Panel A of Figure 4 plots the second-order perturbation approximation for employment,  $n_t$ , when the wage rate is equal to 0.96. It also plots a very accurate solution obtained with an accurate projection method, which we use as a stand-in for the truth. When  $z$  takes on its low value ( $-\zeta$ ), then the perturbation approximation has a second fixed point when  $n$  is (roughly) 0.83. With discrete support this second fixed point is irrelevant in the sense that as long as the economy starts in the ergodic region (also indicated in the figure), then the generated time path is always stable. In particular, even if the economy is extremely unlucky and  $z_t$  is always equal to  $-\zeta$ , then the economy will still converge towards a positive value for  $n_t$ . Similarly, if  $z_t$  is always equal to  $+\zeta$ , then the generated time path will remain well-behaved.

Moreover, the second-order perturbation approximation is a monotone increasing function in the ergodic set. In fact, the second-order perturbation approximation is very close to the very accurate projection approximation in the ergodic set.

Even though the second-order perturbation approximation is doing very well, a small change in the parameter values changes the picture completely. And the properties of the model ask for such a change in the parameter values. As mentioned above, the standard deviation of HP-filtered employment relative to the standard deviation of HP-filtered productivity is still below its empirical counterpart. In Hagedorn and Manovskii (2008), it is shown that this ratio increases when the average surplus decreases. Thus, to get the model properties closer to the observed ones we increase the value of the wage rate,  $w$ , from 0.96 to 0.973.

**second-order perturbation when  $w = 0.973$ .** Panel B of Figure 4 plots the policy function when  $w = 0.973$ . The small increase in the wage rate leads to a minor shift in the policy function. This is true for the "true" policy function and the second-order perturbation approximation. Nevertheless this minor shift has enormous consequences for

---

<sup>23</sup>See Appendix D.1 for a description of the projection method used.

time paths simulated with the second-order perturbation approximation because the time path for employment,  $n_t$ , now explodes.<sup>24</sup> The reason is that the second-order perturbation approximation is now *always* below the 45° line when  $z$  takes on its low value. That is, according to the perturbation approximation there is no longer a bounded ergodic set.

### 3.4 Perturbation approximations and the modified Deaton model

The modified Deaton model resembles the first two models in two aspects. First, it is also a relatively simple model. Second, the agent in the modified Deaton model also faces a trade-off between the return on savings (or cost of borrowing) and consumption smoothing. The model also differs in an important aspect from the first two models. The volatility is much higher at standard parameter values, since this is a model that describes individual behavior and incorporates idiosyncratic uncertainty.

**Parameter values for the modified Deaton model.** We set  $r = 0.03$ ,  $\gamma = 3$ ,  $\bar{z} = 0.4$ ,  $\sigma_z = 0.1$ , and  $\beta = 0.9$ . We choose a low value for  $\beta$  to ensure that agents are sufficiently impatient. If agents are not impatient, then penalty functions or borrowing constraints would not matter. The standard deviation of idiosyncratic income,  $\sigma_z$ , is such that a negative two-standard-deviations shock implies a level of income that is 20% below its mean value. This is clearly not an excessively volatile process for idiosyncratic risk; some papers in the literature even consider specifications for idiosyncratic risk that allow individual income levels to be equal zero.<sup>25</sup> Furthermore, we set  $\eta_0 = 20$ ,  $\eta_1 = 0.04464$ , and  $\eta_2 = 0.00352$ . The values of  $\eta_1$  and  $\eta_2$  are such that the mean and standard deviation of  $a_t$  are equal to the corresponding values in the model with the non-negativity constraint. The value of  $\eta_0$  is modest in the sense that 19% of the times the agent chooses a negative value for  $a_t$ , whereas this cannot happen in the original Deaton model.<sup>26</sup> To get closer to the Deaton model one would need a higher value for  $\eta_0$ , that is, more curvature in the

---

<sup>24</sup>Instability occurs already for lower values of the wage rate, namely when it is close to 0.935.

<sup>25</sup>See, for example Krusell and Smith (1998).

<sup>26</sup>See De Wind (2008) for a more detailed comparison of the model with the inequality constraint and models with penalty functions.

penalty function.

**Instability of second-order perturbation approximation.** Figure 5 plots the "true" policy function and the second-order perturbation approximation.<sup>27</sup> Instead of plotting the savings choice,  $a$ , as a function of the state variable, cash on hand,  $x$ , we plot the expected value of next-period's cash on hand,  $E[x_{+1}|x]$ , which is equal to  $a + E[z_{+1}]$ . The reason is that the graph of the relationship between (the expected value of)  $x_{+1}$  as a function of  $x$  directly reveals whether the dynamics are stable or not.

The true policy function is convex and monotonically increasing for the values of  $x$  observed in the cross-section. Interestingly, the same is true for the second-order perturbation approximation. That is, the second-order perturbation approximation does not exhibit any undesirable shapes in the relevant part of the state space.

The second-order perturbation approximation is unstable, however, because there is a second intersection with the 45°-line above the true steady state. If this intersection is too close to the true steady state, then simulated time paths will eventually take on values to the right of the second intersection at which point the economy is expected to diverge.<sup>28</sup> This turns out to be the case for the parameter values used here.

**Instability of higher-order perturbation approximations.** In Appendix A, we document that the instability problem only gets worse if third-order perturbation is used. The generated time paths are stable for fourth and fifth-order perturbation approximations, but these policy functions display some very large and odd oscillations.

---

<sup>27</sup>The "true" policy function is represented by a very accurate solution obtained with a projection procedure. See Appendix D.2 for details.

<sup>28</sup>Being to the right of the intersection is not sufficient for the asset holdings of the agent to diverge. If the agent receives sufficiently soon sufficiently low values for  $z_t$ , then the asset holdings would drop back into the stable region.

## 4 Stable nonlinear perturbation-based approximations

In this section, we describe two procedures—both based on perturbation analysis—that generate stable time paths. The first procedure solves for the period- $t$  decisions using (i) the exact nonlinear equations of the model for periods  $t$  through  $t+J$  and (ii) the first-order perturbation approximation to describe the behavior in period  $t+J$ . The second procedure is the pruning procedure. The pruning procedure resembles the standard perturbation procedure in that it starts out with a first-order approximation and sequentially adds higher-order monomial terms. In contrast to the standard perturbation approach, the pruning procedure also adds state variables at each step.

### 4.1 Perturbation-plus procedure

To explain the procedure, it is more convenient to write the generic description of the model using only the state variables. That is, we replace Equation (23) with

$$0 = \mathbb{E} [H(x_{+1}, x, x_{-1}, z_{+1}, z)]. \quad (29)$$

The objective is to determine  $x$  given values for  $x_{-1}$  and  $z$ . There are two reasons why Equation (29) is not a standard nonlinear equation in  $x$ . First, the equation contains an integral. Second,  $x_{+1}$  is also unknown. The first issue can be dealt with using numerical integration procedures. To deal with the second issue we replace  $x_{+1}$  by a function of the state variables. If we use the one-step ahead "perturbation-plus" procedure, then  $x_{+1}$  is replaced by the first-order perturbation approximation,  $\tilde{f}_{1st}(x, z_{+1})$ . Thus, we solve  $x$  from

$$0 = \tilde{\mathbb{E}} \left[ H(\tilde{f}_{1st}(x, z_{+1}), x, x_{-1}, z_{+1}, z) \right], \quad (30)$$

where  $\tilde{\mathbb{E}}$  denotes that the integral is calculated using a numerical integration procedure. We denote the value of  $x$  that solves Equation (30) by  $\hat{f}_{+1}(x_{-1}, z)$ , where the subscript indicates that  $x$  is solved using the true model equations for the current period and 1 additional future. There may be no analytical solution for  $\hat{f}_{+1}(\cdot)$  in which case  $\hat{f}_{+1}(\cdot)$  is only implicitly defined by Equation (30). In this case, one has to use a nonlinear equation solver to solve for the value of  $x = \hat{f}_{+1}(x_{-1}, z)$ .

To generate a time path for  $x_t$  one would simply iterate on  $\hat{f}_{+1}(x_{t-1}, z_t)$ . Note that  $\hat{f}_{+1}(x_{t-1}, z_t)$  is constructed under the assumption that next period's value for  $x$  is calculated using  $\tilde{f}_{1st}(x_t, z_{t+1})$ . As next period comes along, however,  $x_{t+1}$  is not calculated using the first-order perturbation approximation,  $\tilde{f}_{1st}(x_t, z_{t+1})$ , but is calculated using  $\hat{f}_{+1}(x_t, z_{t+1})$ .

The two-step ahead modification,  $x = \hat{f}_{+2}(x_{-1}, z)$ , is the value of  $x$  that is the solution to

$$0 = \tilde{\mathbb{E}} \left[ H(\hat{f}_{+1}(x, z_{+1}), x, x_{-1}, z_{+1}, z) \right]. \quad (31)$$

That is, the value of  $x$  is based on (i) the exact equations of the model for this period (ii) the exact equations of the model for the next period, and (iii) the assumption that the behavior in two periods is based on the first-order perturbation approximation.<sup>29</sup>

In theory, one could iterate on this process and construct the  $J$ -step ahead modification,  $x = \hat{f}_{+J}$ . As  $J$  approaches infinity, then  $x$  and  $x_{+1}$  are based on the same policy function, that is, one has a rational expectations equilibrium.

If there is no analytical expression for  $\hat{f}_{+J}(\cdot)$ , then the procedure quickly becomes very expensive as  $J$  increases. Even in a model as simple as the neo-classical growth model, there is no analytical expression for the value of  $x$  that solves Equation (30). As discussed in Appendix C.1, there is a slight modification of the algorithm that is much faster and we found the numerical results to be very similar.<sup>30</sup>

## 4.2 Discussion of the perturbation-plus procedure

The perturbation-plus procedure is quite easy to program. The most tricky part is to implement a numerical integration procedure. The procedure is easier to implement than

---

<sup>29</sup>The perturbation-plus procedure is related to the extended-path method of Fair and Taylor (1983) and Gagnon (1990). This procedure also solves for period- $t$  variables by looking a number of periods into the future. The difference is that the perturbation-plus procedure calculates conditional expectations explicitly, namely by using numerical integration procedures, and it closes the system by using the first-order perturbation solution to describe the *behavior* in the last period, whereas the extended-path method closes the system using terminal conditions for expectations or variables.

<sup>30</sup>This result could very well depend on the problem at hand and not carry over to other models.

projection methods in that it does not require the user to construct a grid. The user also does not have to worry about what class of approximating functions to use (Chebyshev polynomials, splines, etc.). Nevertheless, the procedure has several of the benefits of a projection procedure, at least for sufficiently high  $J$ , because it uses the exact equations of the model and it explicitly approximates the conditional expectation with an accurate numerical integration procedure.

The disadvantage of this procedure is that the problem quickly gets very expensive as  $J$  increases. That is, in practice one can only use this procedure if the appropriate value for  $J$  is low. Whether an accurate solution can be obtained with a low value for  $J$  depends on the model at hand and in particular on the value of  $\beta$ . In Section 5, we discuss an example in which the 1<sup>st</sup> as well as the second-order perturbation approximations are not accurate, but the one-step ahead modification of the first-order perturbation approximation is. But we have also encountered cases where even the thirteen-step ahead modification is not yet accurate (although a lot more accurate than the first-order perturbation approximation itself).

### 4.3 Pruning

In this section, we describe and discuss the procedure proposed in Kim, Kim, Schaumburg, and Sims (2008) to deal with explosive second-order perturbation approximations. A more formal motivation is given in Lombardo (2010) for second-order approximations and in Appendix B for both second and higher-order pruning.

The second-order perturbation approximation can be written as

$$\begin{aligned} & \tilde{f}_{2^{\text{nd}}}(x_{t-1} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z}) \\ & = \\ & \tilde{f}_{2^{\text{nd}}}^{(1)}(x_{t-1} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z}) + \tilde{f}_{2^{\text{nd}}}^{(2)}(x_{t-1} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z}) \end{aligned} \tag{32}$$

where  $\bar{x}_{2^{\text{nd}}}$  is the *stochastic* steady state of the second-order perturbation approximation,  $\tilde{f}_{2^{\text{nd}}}^{(1)}(\cdot)$  is the part of  $\tilde{f}_{2^{\text{nd}}}(\cdot)$  with the linear terms, and  $\tilde{f}_{2^{\text{nd}}}^{(2)}(\cdot)$  is the part of  $\tilde{f}_{2^{\text{nd}}}(\cdot)$  with the second-order terms. The pruning procedure consists of the following steps.

1. Simulate  $x_t^{(1)}$  using

$$x_t^{(1)} - \bar{x}_{2^{\text{nd}}} = \tilde{f}_{2^{\text{nd}}}^{(1)}(x_{t-1}^{(1)} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z}). \quad (33)$$

2. Simulate  $x_t = x_t^{(2)}$  using

$$x_t^{(2)} - \bar{x}_{2^{\text{nd}}} = \tilde{f}_{2^{\text{nd}}}^{(1)}(x_{t-1}^{(2)} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z}) + \tilde{f}_{2^{\text{nd}}}^{(2)}(x_{t-1}^{(1)} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z}), \quad (34)$$

where the values of  $z_t$  used are identical to those used in step 1. The process  $x_t^{(1)} = \tilde{f}_{2^{\text{nd}}}^{(1)}(x_{t-1}^{(1)} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z})$  is stationary unless the Blanchard-Kahn conditions are not satisfied. The process  $\tilde{f}_{2^{\text{nd}}}^{(2)}(x_{t-1}^{(1)} - \bar{x}_{2^{\text{nd}}}, z_t - \bar{z})$  is stationary because both  $x_t^{(1)}$  and  $z_t$  are stationary. Consequently, the simulated values of  $x_t$  are stationary as well.

Give a  $n^{\text{th}}$ -order regular perturbation approximation, the  $n^{\text{th}}$ -order *pruned* perturbation approximation,  $x_t = x_t^{(n)}$ , is generated using the following iterative scheme:<sup>31,32</sup>

$$x_t^{(j)} - \bar{x}_{J^{\text{th}}} = \sum_{i=1}^j \tilde{f}_{J^{\text{th}}}^{(i)}(x_{t-1}^{(j-i+1)} - \bar{x}_{J^{\text{th}}}, z_t - \bar{z}; \sigma) \text{ for } j = 1, \dots, n. \quad (35)$$

For  $n \geq 2$ , there is an alternative. First calculate  $x_t^{(1)}$  and  $x_t^{(2)}$  using Equation (35). Next calculate the  $n^{\text{th}}$ -order approximation using

$$x_t^{(j)} - \bar{x}_{J^{\text{th}}} = \tilde{f}_{J^{\text{th}}}^{(1)}(x_{t-1}^{(j)} - \bar{x}_{J^{\text{th}}}, z_t - \bar{z}; \sigma) + \sum_{i=2}^j \tilde{f}_{J^{\text{th}}}^{(i)}(x_{t-1}^{(j-i)} - \bar{x}_{J^{\text{th}}}, z_t - \bar{z}; \sigma) \text{ for } j = 3, \dots, n. \quad (36)$$

The formulation in Equation (36) uses the highest available (stationary) variable. The specification in Equation (35) is also a  $n^{\text{th}}$ -order approximation since  $x_t^{(j)}$  is used in  $(n - j + 1)^{\text{th}}$ -order monomials.

---

<sup>31</sup>For third and higher-order perturbation solutions it is in theory possible that the Blanchard-Kahn conditions are (just) satisfied and the first-order perturbation solution is stable, but that the correction for  $\sigma$  makes the linear part of the third and higher-order perturbation solutions unstable. We suspect that such knife-edge cases are rare.

<sup>32</sup>For second-order perturbation, the value of  $\sigma$  at most affects the stochastic steady state, but for higher-order perturbation the value of  $\sigma$  could affect other coefficients. For this reason we introduce  $\sigma$  again as an argument of the perturbation approximation.

There is a slight difference between the procedure described here and the way pruning is implemented in Kim, Kim, Schaumburg, and Sims (2008) and Lombardo (2010). In each step of the  $n^{\text{th}}$ -order pruned perturbation procedure, we use the coefficients of the  $n^{\text{th}}$ -order perturbation approximation. For example, to generate  $x_t^{(1)}$  we use the linear part of the  $n^{\text{th}}$ -order perturbation approximation whereas Kim, Kim, Schaumburg, and Sims (2008) and Lombardo (2010) use the first-order perturbation approximation. As is documented in Appendix B, both approximations are  $n^{\text{th}}$ -order approximations. But our formulation has an important practical advantage.

The advantage of our approach is that the different measures of the state variable, i.e.,  $x_t^{(1)}$ ,  $x_t^{(2)}$ , etc., have the same steady state value, namely the stochastic steady state according to the  $n^{\text{th}}$ -order perturbation approximation. For example, suppose that the stochastic steady state is substantially different from the non-stochastic steady state according to the second-order perturbation approximation—a situation one can easily encounter when volatility is so high that second-order perturbation approximations do not generate stable time paths. If the procedure of Kim, Kim, Schaumburg, and Sims (2008) and Lombardo (2010) is used to obtain higher-order pruned perturbation approximations, then the law of motion for  $x_t^{(1)}$  would be the same independent of the order of the approximation considered. This means that any systematic bias would not disappear either. In contrast, according to our pruning procedure the law of motion for  $x_t^{(1)}$  would change if higher-order pruned perturbation approximations are considered.<sup>33</sup>

#### 4.4 Discussion of pruning

Given the regular perturbation approximation, it is straightforward to obtain the pruned modification. Given that it always generates stable time paths, it is definitely an approach that should be taken seriously. However, there are several undesirable features of the pruning approximation. The user should at least be aware of these features and ask him/herself whether these are important for the problem at hand.

---

<sup>33</sup>As discussed in Appendix B, both procedures generate valid  $n^{\text{th}}$ -order perturbation approximations. The reason is that the stochastic steady state and the non-stochastic steady state converge towards each other as  $\sigma \rightarrow 0$ .



#### 4.4.1 Inability of pruning to fit standard polynomials

Pruned perturbation approximations as well as regular perturbation approximations are polynomials. If the truth is an  $n^{\text{th}}$ -order polynomial, then the *regular*  $n^{\text{th}}$ -order perturbation approximation would give an exact fit. But the pruned  $n^{\text{th}}$ -order perturbation approximation is not able to accomplish this. We use the following example to explain the reason.

Suppose that the truth is given by

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-1}^2 + \sigma \varepsilon_t \quad (37)$$

with

$$\mathbb{E}_t [\varepsilon_t^2] = 1.$$

The second-order pruned perturbation approximation is generated by the following system:

$$x_t^{(1)} = \alpha_1 x_{t-1}^{(1)} + \sigma \varepsilon_t, \quad (38)$$

$$x_t^{(2)} = \alpha_1 x_{t-1}^{(2)} + \alpha_2 \left( x_{t-1}^{(1)} \right)^2 + \sigma \varepsilon_t. \quad (39)$$

Although the law of motion for  $x_t$ , given in Equation (39), closely resembles the true law of motion, given in Equation (37), there is one fundamental difference. The difference is that the expression in Equation (39) contains  $x_{t-1}^{(1)}$ , which is generated by a different law of motion than  $x_t$ .

Now consider higher-order approximations. When the truth is given by Equation (37), then the third-order *regular* perturbation approximation is, of course, equal to the second-order regular perturbation approximation, which in turn is equal to the truth. But this is not true for the pruned perturbation approximation. The third-order pruned perturbation approximation is generated by the following set of equations:

$$x_t^{(1)} = \alpha_1 x_{t-1}^{(1)} + \sigma \varepsilon_t, \quad (40)$$

$$x_t^{(2)} = \alpha_1 x_{t-1}^{(2)} + \alpha_2 \left( x_{t-1}^{(1)} \right)^2 + \sigma \varepsilon_t, \quad (41)$$

$$x_t^{(3)} = \alpha_1 x_{t-1}^{(3)} + \alpha_2 \left( x_{t-1}^{(2)} \right)^2 + \sigma \varepsilon_t. \quad (42)$$

The third-order pruned perturbation approximation still does not match the true second-order polynomial, although the mistake has become smaller.<sup>34</sup>

There is a neat way to characterize the error made by the pruned perturbation approximation. Suppose one uses the pruned perturbation approximation to calculate the impulse response function (IRF) of a unit-shock to  $\varepsilon_t$  starting at the steady state. If the truth is a second-order polynomial, then the  $n^{\text{th}}$ -order pruned perturbation approximation will give the right values for the IRF for the first  $n$  periods. Thus, the second-order pruned perturbation approximation will give the right answer only up to the first two periods.

The policy functions of many economic models can be approximated well with a low-order polynomial. In those cases, regular perturbation would make more sense than pruned perturbation. If the time paths generated with a regular perturbation approximation explode, then such low-order perturbation approximations are apparently not accurate. Whether pruned perturbation approximations are accurate in such cases remains an open question on which we will shed some light in Section 5.

#### 4.4.2 Pruning approximation is not a function of the original state variables

In the DSGE literature, it is common to focus on models of which the solutions are recursive functions of a set of predetermined endogenous state variables and exogenous state variables (here  $x_{-1}$  and  $z$ ). For many models one can prove that such a recursive equilibrium exists. Numerical approximations are almost always a function of the same set of state variables, although the effect of some state variables may be set to zero or turn out to be approximately equal to zero in the approximation. The pruned perturbation approximation is also a recursive function, but of a larger (and potentially much larger) set of variables.<sup>35</sup> It is a *correspondence* in the original state variables,  $x_{-1}$  and  $z$ .

To illustrate this aspect of the pruned perturbation approximation, we use the neo-

---

<sup>34</sup>The variable that is being squared is now being generated by a law of motion that is closer to the truth.

<sup>35</sup>The pruned perturbation approximation, like the true solution, is also a function of the values of the endogenous state variables in the first period and the *complete* history of the exogenous state variables.

classical growth model.<sup>36</sup> We assume that the law of motion for productivity,  $z_t$ , is equal to the first-order Markov process described in Section 2.2. The advantage of this process is that  $z_t$  takes on only two values, which makes it possible to graphically document whether the policy function is close to a function of the original state variables or not.

To see whether the pruned perturbation approximation is close to a function of the original state variables we do the following. Using the pruned second-order perturbation approximation, we generate a long time series for capital,  $k_t^{(2)}$ .<sup>37</sup> Next, we plot the change in capital,  $k_t^{(2)} - k_{t-1}^{(2)}$ , as a function of the state variable,  $k_{t-1}^{(2)}$ . One set of numbers will be for the high value of  $z_t$  and one set for the low value of  $z_t$ . The results are reported in Figure 6. The figure clearly illustrates that the pruned perturbation approximation does not come close to being a function of the original state variables. The value of  $\zeta$  used to generate the graph is equal to 0.6405, which means that the amount of uncertainty is high for a macroeconomic model. But it does not make sense to analyze this issue for standard parameter values, since the solutions to this very simple model are then approximated well with first-order perturbation and there would be no reason to consider higher-order pruned perturbation.

The cause of this problematic aspect of the pruned perturbation approximation is that additional state variables are added as arguments of the approximation. If the behavior of the additional state variable,  $k_{t-1}^{(1)}$  is similar to the behavior of the actual state variable,  $k_{t-1}^{(2)}$ , then the dimension of the state space only increases nominally, but is of little importance. This way of characterizing the problem is useful for the discussion below. In Figure 7, we present, therefore, a scatter plot of simulated values for the two measures for capital generated by the pruned second-order perturbation approximation,  $k_t^{(2)}$  and  $k_t^{(1)}$ . The values of  $k_t^{(2)}$  and  $k_t^{(1)}$  are equal to each other in the first observation, but the pair immediately leaves the 45° line. The graph makes clear that these two state variables behave very differently. Not only do they have different means, the thickness of the scatter plot

---

<sup>36</sup>Den Haan and De Wind (2009) illustrate this property of pruned perturbation approximations using the modified Deaton model.

<sup>37</sup>The simulation is based on the following parameter values:  $\alpha = 0.36$ ,  $\beta = 0.99$ ,  $\gamma = 3$ ,  $\delta = 0.025$ ,  $\bar{\rho}_z = 0.975$ , and  $\zeta = 0.6405$ .

makes clear that the correlation is not that high. Moreover,  $k_t^{(1)}$  even takes on negative values, whereas this is never the case for  $k_t^{(2)}$ .

#### 4.4.3 Why should approximations be functions of the original state variables?

In this subsection, we give reasons why it is problematic if a numerical approximation is *not* a function of the original state variables. One could argue that a numerical approximation always makes mistakes in some dimensions. In particular, if one is only interested in generating time paths for the model variables, then the issues we raise here do not directly show up. Indirectly, they could show up. If the numerical approximation is not close to being a function of the original state variables, then the approximation is inaccurate in at least some aspects and these inaccuracies are likely to show up in the simulated data as well.<sup>38</sup>

**Properties of the policy function.** Two typical reasons for calculating a numerical solution are to study the properties of the policy function and to analyze how these properties depend on the values of the structural parameters. For example, one may be interested to know how the marginal propensity to consume depends on the wealth level and how this dependence changes with key parameters like the coefficient of relative risk aversion or the amount of uncertainty the agent faces. When using a pruned perturbation approximation, there are multiple measures for each state variable. And the higher the order of the approximation, the more measures. Moreover, the results in Figure 7 make clear that the different measures of the same state variables can behave quite differently. This means that it would be inconsistent with the proposed numerical approximation to simply ignore the differences between the various measures of the state variables.

Of course, one could limit oneself to study the function at the steady state—where the alternative measures take on identical values—but this only provides very limited insights into the properties of the policy function and would rule out analyzing the effect of changes in  $\sigma$ .

---

<sup>38</sup>The accuracy of simulated time paths is discussed in Section 5.

**Properties of the impulse response functions.** Properties of the model are often characterized by the model's IRFs. If the approximation is nonlinear, then the IRF depends on the values the state variables take on when the shock occurs. Suppose one wants to know how different the IRFs of the matching model are when the shock occurs in a tight labor market (high value of  $n_{t-1}$ ) and in a slack labor market (low value of  $n_{t-1}$ ). If the numerical solution is a function of  $n_{t-1}$ , then one could simply choose two representative values for  $n_{t-1}$ . When using second-order *pruned* perturbation approximations one also has to choose the corresponding representative values for  $n_{t-1}^{(1)}$ . For higher-order *pruned* perturbation approximations, one has to figure out the values of even more measures of the state variables. The results in Figure 7 make clear that there is a systematic difference between the first-order measure of the state variable and the second-order measure of the same state variable. This means that it would be misleading to simply set  $n_{t-1}^{(1)}$  equal to  $n_{t-1}$ .

**Forecasting.** If one wants to use the model to forecast, then one has to know the values of the state variables. For example, suppose one wants to forecast using the matching model. If one sets  $n_{t-1}^{(1)}$  equal to the observed employment rate, then one makes an assumption that is inconsistent with the numerical approximation being used, namely one assumes that  $n_{t-1} = n_{t-1}^{(1)}$ , whereas we have shown that this is not true.

**Incorporating a numerical solution into a different numerical problem.** Numerical approximations are also used as inputs in other (numerical) problems. For example, one may first solve for the numerical approximation of the private sector and then determine government behavior. The policy rule describing government behavior can then be an input in the problem that determines the behavior of the private sector and one would iterate until the problem has converged. Another example can be found in Den Haan and Rendahl (2010) who use the numerical solution for individual behavior to derive the behavior of the corresponding aggregates by explicitly aggregating.

It may not be impossible to do these types of exercises with *pruned* perturbation approximations, but the additional state variables introduced by the pruning procedure

would make these exercises more cumbersome.

## 5 Accuracy of the stable perturbation-based procedures

In this section, we evaluate the accuracy of the pruned perturbation approximations and the perturbation-plus approximations for the three models of Section 2. For the Brock-Mirman model, we compare a time path of the approximation with the corresponding time path of the true solution. For the other models, we use a time path generated by a very accurate projection method instead of the (unknown) true solution.<sup>39</sup> The length of the time path,  $T$ , is set equal to 10,000.

The distance between a period- $t$  variable generated with an approximation and the corresponding "true" value is typically measured as the absolute percentage error. This will not make sense if—as is the case in several of our models—variables take on values that are close to zero. An alternative would be to scale the (absolute) difference between the period- $t$  value according to the approximation and the corresponding true value by the time series average of the true series. But this measure would not make sense if variables take on values that are bigger than multiple times this mean and we encounter observations where the true value is more than 20 times the mean. For those type of observations, the regular percentage error would be more appropriate.

Therefore, we define the period- $t$  error as

$$e_t = \min \left\{ \left| \frac{\tilde{x}_t - x_t}{x_t} \right|, \left| \frac{\tilde{x}_t - x_t}{\bar{x}_T} \right| \right\}, \quad (43)$$

where  $\tilde{x}_t$  is the realization according to the approximation,  $x_t$  the realization according to the true (or very accurate projection method) solution, and  $\bar{x}_T$  is the mean value of  $x_t$ . We report both the maximum and the mean of  $e_t$ . For all three models, we report only the results for the state variable. The errors for consumption are somewhat smaller, but the conclusions drawn here do not depend on the variable considered.

---

<sup>39</sup>That the projection method generates very accurate solutions is established in Appendix D.

### 5.1 Accuracy of approximations to the Brock-Mirman model

We consider two values for  $\sigma_z$ , namely  $\sigma_z = 0.1$  and  $\sigma_z = 0.2$ . Table 1 reports information about the approximation errors and Table 2 reports summary statistics for the behavior of capital according to the different policy rules. The following observations can be made.

First, the first-order perturbation approximation performs very poorly. This approximation generates large *negative* values for the capital stock. Moreover, the maximum value generated by the first-order perturbation approximations is less than half of the true maximum value when  $\sigma_z = 0.1$  and less than 13% of the true maximum when  $\sigma_z = 0.2$ . Not surprisingly, the accuracy measures are very poor.

Second, the pruned second-order approximation performs worse than the regular second-order perturbation approximation. Although, the pruned second-order approximation typically does substantially better than first-order perturbation, its maximum error is higher when  $\sigma_z = 0.20$ . Whereas the problem of first-order perturbation is that it generates values for capital that are way too low, the problem for both types of second-order perturbation approximations is that the lowest values are substantially above the true minimum value. The gap is roughly equal to one quarter of the mean capital stock.

Third, our proposed alternative does substantially better than the second-order perturbation approximations, both in terms of having better accuracy measures and in terms of generating summary statistics that are closer to the truth. Nevertheless, the maximum errors are non-trivial. For the two-step ahead perturbation-plus approximation, the maximum errors are equal to 5.2% and 8.1% when  $\sigma_z$  is equal to 0.1 and 0.2, respectively. This is not spectacular, but substantially better than the second-order pruned approximation for which the corresponding numbers are 47.9% and 193.8%.

Figure 8 plots that part of the sample where the largest errors are obtained by the perturbation-plus approximations when  $\sigma_z = 0.2$ . This truly is an usual period: the true value of  $k_t$  takes on a value that is more than twenty times the sample average. The two-step ahead perturbation-plus approximation does a good job following the true time path, but doesn't reach the same peak. It reaches a maximum value of 6.3 whereas the true maximum is equal to 6.9. The maximum reached by second-order pruned perturbation is

only 2.2.

The three-step ahead perturbation-plus approximation does even better than the two-step ahead approximation. For example, the maximum reached during this enormous upswing is equal to 6.68, only 2.7% below the true maximum (which is 2000% above the average value). The problem is that the three-step ahead perturbation-plus approximation is quite expensive to run.

## 5.2 Accuracy of approximations to the matching model

As in Section 3.3, we consider two values for the wage rate, namely  $w = 0.96$  and  $w = 0.973$ . At the lower value of  $w$  the regular second-order perturbation approximation is stable and at the higher value of  $w$  it is not. Accuracy measures and summary statistics for the different numerical approximations are given in Tables 1 and 2, respectively.<sup>40</sup> The following observations can be made.

First, the error measures are much smaller than for the previous model. The reason for this is the much smaller variability. By construction, the model generates plausible employment volatility (relative to labor productivity) when  $w = 0.973$ . In particular, the employment rate fluctuates between 89.8% and 96.4% according to the "true" solution. The model generates too little volatility in the employment rate when  $w = 0.96$ .

Second, first-order perturbation again performs very poorly. When  $w = 0.973$ , the first-order perturbation approximation predicts that the mean employment rate is equal to 94.9% while in fact it should be 93.1%. Making a mistake of 1.8 percentage point when the range of generated values is only 6.6 percentage point is very troubling.

Third, when  $w = 0.973$  the regular second-order perturbation approximation explodes and is obviously outperformed by pruned perturbation. When  $w = 0.96$  and the regular second-order *just* does not explode, it outperforms pruned perturbation by far. In fact, it

---

<sup>40</sup>The minimum values obtained by the second-order perturbation approximations are the same when  $\sigma = 0.1$  and when  $\sigma = 0.2$ . The reason is the following. First, the correction term for uncertainty is almost zero. Second, when  $z$  is low then the (quadratic) policy function is (i) close to the minimum and (ii) quite flat. This means that the chosen values of  $k$  hover around the same value until a sufficiently high value for  $z$  pushes the economy out of this area.



performs very well.

Fourth, the results for second-order pruned perturbation are substantially better than those for first-order perturbation. Nevertheless, the time paths generated by second-order pruned perturbation are inaccurate in some important dimensions. In particular, it misses the potential depth of recessions. Consider the case when  $w = 0.973$ . The worst that can happen according to the second-order pruned perturbation approximation is a drop in the employment rate to 91.4%, while the actual minimum is 89.8%. In contrast, second-order pruned perturbation predicts the maximum value quite well. It predicts a peak of 96.35% whereas our accurate projection method predicts a peak of 96.44%. Consistent with these observations, second-order pruned perturbation underestimates the standard deviation of employment with 35%.

Fifth, the perturbation-plus procedure does very poorly unless the number of forward looking steps is very high. For the Brock-Mirman model, we found that even the one-step ahead perturbation-plus procedure delivered a substantial improvement over first-order perturbation. For this model, the one-step ahead perturbation-plus procedure leads to only a small improvement. As documented in the Table, even the nine-step ahead perturbation-plus procedure performs worse than the second-order pruned perturbation approximations. To get performance that is comparable to that of the second-order pruned perturbation approximation, the number of forward looking steps has to be at least 15 steps.<sup>41</sup>

If we replace the first-order perturbation approximation by *two* linear functions of  $n_{-1}$ , namely one for each of the two values of  $z$ , then the perturbation-plus procedure does much better. In fact, even the one-step ahead perturbation-plus procedure outperforms

---

<sup>41</sup>We also considered some cases when  $z$  had continuous support. Interestingly, one-step ahead perturbation plus then did substantially improve upon first-order perturbation and also did much better than the two second-order perturbation approximations. The advantage of discrete support for second-order perturbation approximations is that it limits the maximum deviations from the steady state, which is useful given the nonlinear dependence of  $n$  on  $z$  is not captured well with a second-order perturbation approximation.

the second-order pruned perturbation by far. The two linear functions were obtained with a projection method. It is not difficult to obtain these two linear functions and one always achieves a stunning improvement as long as one conditions on the value of  $z$ . But we do not want to pursue this modification, since the whole idea of the perturbation-plus procedure is that one does not have to worry about setting up a grid.

### 5.3 Accuracy of approximations to the modified Deaton model

As documented in Table 1, there are substantial differences between the time paths generated by the approximations and the time path generated by the accurate projection method. For example, the maximum error for the first-order perturbation approximation is 137% and the average error is equal to 44%. This is not a trivial model to solve, given that the amount of idiosyncratic risk is substantial. For example, according to the accurate projection approximation, asset holdings range between  $-0.094$  and  $0.848$ , compared with a mean of  $0.085$ .

Substantial improvements are obtained by the nonlinear approximations, except by the regular second-order perturbation approximation, because it generates a time path that is not stable. The maximum error of the second-order pruned perturbation approximation is equal to 127%, only a slight improvement over the one obtained by the first-order perturbation approximation. But the average error is substantially less, namely 12.2%. The average error of the two-step ahead perturbation-plus procedure is slightly smaller than the one for second-order pruning, but the maximum error is substantially smaller, namely 64.6%.

Table 2 documents that there are also substantial differences in the properties of the generated time paths. The two-step ahead perturbation-plus procedure predicts a standard deviation that is 10% below the true value and second-order pruning underpredicts this standard deviation with 16%. The approximations have most difficulty in following the true time path in extreme situations. Interestingly, the perturbation-plus procedure has problems with the peaks and pruning has problems with the troughs.

Panel A of Figure 9 plots that part of the generated time path where the perturbation-

plus procedure makes the biggest error and Panel B plots that part where pruning makes the biggest error. First consider Panel A. The figure makes clear that the perturbation-plus procedure is not as bad as the results reported above indicate. At the peak, the two-step ahead perturbation-plus approximation underestimates the maximum achieved level of asset holdings with 16%. Given that the true maximum is ten times as big as the mean asset holdings, this error is not that worrisome. In particular, the perturbation-plus approximation also predicts an enormous increase and some time before and some time after this unusual situation the perturbation-plus procedure does track the accurate time path reasonably well. The maximum error for the two-step ahead perturbation-plus approximation is equal to 64%. Interestingly, the maximum error is not obtained at the peak but after the fall back towards normal levels.<sup>42</sup>

Now turn to Panel B of Figure 9 that plots that part of the time path during which the second-order pruning procedure makes the biggest error. Interestingly, this occurs during a serious downturn. Again, the problem is concentrated in the trough; before and after the trough the pruned perturbation approximation follows the accurate time path quite closely. The error made by second-order pruning highlighted in Panel B may be a bit more severe than the error made by the perturbation-plus approximation in the top panel. First, the trough highlighted in Panel B is much closer to the mean than the peak in Panel A. Moreover, the generated patterns is actually a bit different as well. After the time path generated with pruned perturbation has reached its minimum value, it starts a strong upward movement, whereas the accurate time path continues to drop.

## 6 Concluding comments

In this paper, we have focused on non-trivial numerical problems. In particular, we have looked at models for which regular second-order perturbation generates time paths that reach that part of the state space where the derivative of the second-order approximation has the wrong sign and we have looked at models where the parameter values are just below the critical levels that ensure stability or are just above these critical values.

---

<sup>42</sup>The error made by second-order pruned perturbation is actually bigger at this point.

These are models for which linear approximations do not suffice. The two perturbation-based alternatives to regular perturbation approximations generate stable time paths and incorporate nonlinearities. These methods can deliver substantial improvements over linear approximations. In the examples considered in this paper, the proposed alternatives provide a good *qualitative* insight in how the true time path differs from the path generated by first-order perturbation. In most cases, however, there are important quantitative differences between the time paths generated by the approximations and the accurate solution method. The two exceptions are the three-step ahead perturbation-plus procedure when applied to the Brock-Mirman and the modified Deaton model. But this procedure is expensive to run when three steps of forward looking behavior are used.

Thus, if the user is interested in precise quantitative properties of the model, then it is important to evaluate the quality of the approximation with an accuracy test. It is always a good idea to do this, but this is especially the case when solving models that are such that regular low-order perturbation approximations generate unstable time paths.

## A Higher-order perturbation

The question arises whether one can avoid or at least reduce the severity of the instability of the second-order perturbation approximation by going to higher-order approximations. To shed light on this question we plot the third, fourth, and fifth-order perturbation approximations for the modified Deaton model in Figure 10.

The figure documents that the third-order perturbation approximation also has a second intersection with the 45°-line and that this undesirable intersection is *closer* to the steady state than the undesirable additional intersection for the second-order perturbation approximation. Consistent with this fact, we find that series generated with the third-order perturbation approximation start an explosive trajectory faster than the series generated with the second-order perturbation approximation.

The fourth and fifth-order perturbation approximation are stable, but the stability comes at the cost of having a sharply decreasing policy function in some part of the state space. The histogram below the graph displays the observed distribution of the state variable (according to the accurate projection approximation). It documents that the decreasing part occurs in a very relevant part of the state space for the fifth-order approximation.

## B More on pruning

Kim, Kim, Schaumburg, and Sims (2008) and Lombardo (2010) describe how to construct a second-order pruned perturbation approximation. Lombardo (2010) also explains why his formulation of the second-order pruned perturbation approximation is a second-order local approximation. The formulation in Lombardo (2010) is somewhat different than the one used in Kim, Kim, Schaumburg, and Sims (2008) and is somewhat different than the one used in this paper.

This appendix has three objectives. First, we want to discuss the similarities and differences between the formulation used in Lombardo (2010) and the one used in this paper. Second, we want to explain why a second-order pruned perturbation approxima-

tion is a second-order approximation even though the second-order pruned perturbation approximation of a second-order polynomial is not simply that second-order polynomial. The third purpose of this appendix is to explain how to and how *not* to do higher-order pruning.

## B.1 Different formulations for pruning

The differences are explained using a simple example. Throughout this subsection, we assume that the true law of motion is given by

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-1}^2 + \sigma \varepsilon_t, \quad (44)$$

$$\mathbb{E}[\varepsilon_t^2] = 1. \quad (45)$$

We assume that  $|\rho_1| < 1$ , which implies that the process is locally stable. The regular second-order perturbation approximation is simply equal to the function itself, that is,

$$x_t = \tilde{f}_{2^{\text{nd}}}(x_{-1}; \sigma) = \rho_1 x_{t-1} + \rho_2 x_{t-1}^2 + \sigma \varepsilon_t. \quad (46)$$

Lombardo (2010) generates the second-order pruned perturbation approximation using the following system of equations:

$$\tilde{x}_t^{(2)} = \sigma \tilde{x}_t^{[1]} + \sigma^2 \tilde{x}_t^{[2]}, \quad (47a)$$

$$\tilde{x}_t^{[1]} = \rho_1 \tilde{x}_{t-1}^{[1]} + \varepsilon_t, \quad (47b)$$

$$\tilde{x}_t^{[2]} = \rho_1 \tilde{x}_{t-1}^{[2]} + \rho_2 \left( \tilde{x}_{t-1}^{[1]} \right)^2. \quad (47c)$$

We generate the second-order pruned perturbation approximation using

$$\hat{x}_t^{(2)} = \rho_1 \hat{x}_{t-1}^{(2)} + \rho_2 \left( \hat{x}_{t-1}^{(1)} \right)^2 + \sigma \varepsilon_t, \quad (48a)$$

$$\hat{x}_t^{(1)} = \rho_1 \hat{x}_{t-1}^{(1)} + \sigma \varepsilon_t. \quad (48b)$$

Suppose that

$$\tilde{x}_0^{(1)} = \tilde{x}_0^{(2)} = \hat{x}_0^{(1)} = \hat{x}_0 = x_0 = 0 \quad (49)$$

and

$$\varepsilon_t = 1 \text{ if } t = 1, \quad (50)$$

$$\varepsilon_t = 0 \text{ if } t > 1. \quad (51)$$

For this set of values for  $\varepsilon_t$ ,<sup>43</sup>

$$\tilde{x}_t^{(2)} = \hat{x}_t^{(2)} \neq x_t \text{ and} \quad (52)$$

$$\tilde{x}_t^{(2)} - x_t = \hat{x}_t^{(2)} - x_t = O(\sigma^3). \quad (53)$$

Nevertheless, there is a difference between the two formulations. This is easy to see when  $\sigma = 0$ . The formulation according to Equation (47) implies that

$$\tilde{x}_t^{(2)} = 0 \quad \forall t, \quad (54)$$

whereas the formulation according to Equation (48) implies that

$$\hat{x}_t^{(2)} = \rho_1 \hat{x}_{t-1}^{(2)} + \rho_2 \left( \hat{x}_{t-1}^{(1)} \right)^2, \quad (55a)$$

$$\hat{x}_t^{(1)} = \rho_1 \hat{x}_{t-1}^{(1)}. \quad (55b)$$

That is, the formulation of Lombardo (2010) does not describe any transition dynamics, whereas our formulation does. Nevertheless, it is easy to show that both formulations are proper second-order approximations *if* an additional condition is satisfied. Suppose that  $x_t$  is generated by Equation (44),  $\tilde{x}_t^{(2)}$  is generated by Equation (47), and  $\hat{x}_t^{(2)}$  is generated by Equation (48). Then it is easy to show that

$$\tilde{x}_t^{(2)} = x_t + O(\sigma^3) \text{ and} \quad (56)$$

$$\hat{x}_t^{(2)} = x_t + O(\sigma^3) \quad (57)$$

if

$$x_0 = \sigma^2 \bar{x}_0 \text{ with } \bar{x}_0 < \infty. \quad (58)$$

That is, our pruned perturbation formulation does allow transition dynamics, but the initial value chosen cannot be too far away from the steady state. Similarly, it is acceptable to ignore transition dynamics, but only if the initial value is close enough to the steady

---

<sup>43</sup>A function  $f(\sigma) = O(\sigma^m)$  if there exist an  $M$  such that

$$\lim_{\sigma \rightarrow 0} \frac{f(\sigma)}{\sigma^m} < M < \infty.$$

state. It is obvious that a condition like this is needed. There are values for  $x_0$  such that the time path generated by the true law of motion given in Equation (44) explodes, whereas the time paths generated by Equations (47) and (48) never explode.

## B.2 Pruned perturbation and inability to fit polynomials exactly

Suppose that  $x_t$  is generated by Equation (44) with  $\sigma = 0$ . Thus

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-1}^2. \quad (59)$$

The idea behind perturbation is that the approximation becomes arbitrarily close to the truth if the volatility goes to zero. If  $x_t$  is generated by the law of motion given in Equation (59), then the value of  $\sigma$  has already approached zero. In addition, the truth is a simple second-order polynomial. Nevertheless, neither the pruning approximation according to Equation (47),  $\hat{x}_t^{(2)}$ , nor the pruning approximation according to Equation (48),  $\hat{x}_t^{(2)}$ , provides an exact fit. These pruned perturbation approximations are still proper second-order approximations, but one does need to assume that the initial condition is sufficiently close to the steady state. More formally, even though the true value of  $x_t$  is a standard second-order polynomial of only  $x_{t-1}$  and does not depend on a shock, the second-order *pruned* perturbation approximations are only  $O(\sigma^3)$ , where  $\sigma$  now refers to the distance of the initial condition to the steady state as in Equation (58). In contrast, the regular second-order perturbation approximation is exact and is, thus, of order  $O(\sigma^m)$  for any  $m$ .

It is easy to construct examples in which the dynamics predicted by Equation (59) are quite different from the dynamics predicted by the pruned perturbation approximations even when the dynamics according to the true law of motion are stable.<sup>44</sup> The question then arises whether it is desirable that pruned perturbation approximations do not (approximately) replicate a polynomial when the truth is (approximately) a polynomial. Many functions in economic models can be approximated well with a low-order polynomial

---

<sup>44</sup>For example, suppose that  $x_0 = 1$ ,  $\rho_1 = 0.01$ , and  $\rho_2 = 0.98$ . The response according to the pruned approximation is roughly zero after two periods. In contrast, according to the true law of motion the value of  $x_t$  declines gradually towards zero and gets close to zero only after nine periods.



of the original state variables. The distortion that is introduced by the pruned perturbation approximations is in those cases likely to deteriorate the approximation. But not all functions are approximated well with a polynomial. That is likely to be the case when regular perturbation approximations explode and the true model solution does not. But the fact that regular polynomials provide a poor approximation does, of course, not imply that pruned perturbation approximations do. There are many aspects to a function and stability is only one of them.

### B.3 Convergence of our pruning formulation

In this subsection, we discuss in more detail why our formulation of  $n^{\text{th}}$ -order pruned perturbation generates approximations that are of order  $O(\sigma^{n+1})$ .

#### B.3.1 Convergence of second-order pruning

To simplify the notation, we assume that the true law of motion is defined by

$$0 = E_t[H(x_{t+1}, x_t, x_{t-1}, z_t; \sigma)], \quad (60)$$

$$z_t = \sigma \varepsilon_t, \quad (61)$$

where  $x_t$  and  $\varepsilon_t$  are scalars. The regular second-order perturbation approximation can be written as

$$X_{t;\sigma}^{(2)} = \bar{f}'_x X_{t-1;\sigma}^{(2)} + \bar{f}'_z \sigma \varepsilon_t + 0.5 \bar{f}''_{xx} \left( X_{t-1;\sigma}^{(2)} \right)^2 + \bar{f}''_{xz} X_{t-1;\sigma}^{(2)} \sigma \varepsilon_t + 0.5 \bar{f}''_{zz} (\sigma \varepsilon_t)^2, \quad (62a)$$

$$x_{t;\sigma}^{(2)} = X_{t;\sigma}^{(2)} + \bar{x}_{2^{\text{nd}};\sigma}. \quad (62b)$$

Note that  $X_{t;\sigma}^{(2)}$  is defined as the value  $x_t$  relative to the *stochastic* steady state of the second-order approximation, not relative to the non-stochastic steady state. This means that its value depends on  $\sigma$ . Our formulation of the pruned second-order perturbation

approximation,  $\hat{X}_{t;\sigma}^{(2)}$  is given by

$$\hat{X}_{t;\sigma}^{(1)} = \bar{f}'_x \hat{X}_{t-1;\sigma}^{(1)} + \bar{f}'_z \sigma \varepsilon_t, \quad (63a)$$

$$\begin{aligned} \hat{X}_{t;\sigma}^{(2)} = & \bar{f}'_x \hat{X}_{t-1;\sigma}^{(2)} + \bar{f}'_z \sigma \varepsilon_t \\ & + 0.5 \bar{f}''_{x^2} \left( \hat{X}_{t-1;\sigma}^{(1)} \right)^2 + \bar{f}''_{xz} \hat{X}_{t-1;\sigma}^{(1)} \sigma \varepsilon_t + 0.5 \bar{f}''_{z^2} (\sigma \varepsilon_t)^2, \end{aligned} \quad (63b)$$

$$\hat{x}_{t;\sigma}^{(2)} = \hat{X}_{t;\sigma}^{(2)} + \bar{x}_{2^{\text{nd}};\sigma}. \quad (63c)$$

The formulation for the pruned perturbation approximation used in the literature is given by

$$\tilde{X}_t^{(1)} = \bar{f}'_x \tilde{X}_{t-1}^{(1)} + \bar{f}'_z \sigma \varepsilon_t, \quad (64a)$$

$$\begin{aligned} \tilde{X}_{t;\sigma}^{(2)} = & \bar{c}_{2^{\text{nd}};\sigma} + \bar{f}'_x \tilde{X}_{t-1}^{(2)} + \bar{f}'_z \sigma \varepsilon_t \\ & + 0.5 \bar{f}''_{x^2} \left( \tilde{X}_{t-1}^{(1)} \right)^2 + \bar{f}''_{xz} \tilde{X}_{t-1}^{(1)} \sigma \varepsilon_t + 0.5 \bar{f}''_{z^2} (\sigma \varepsilon_t)^2, \end{aligned} \quad (64b)$$

$$\hat{x}_{t;\sigma}^{(2)} = \tilde{X}_{t;\sigma}^{(2)} + \bar{x}, \quad (64c)$$

where  $\bar{x}$  is the non-stochastic steady state.

Our formulation ensures that the stochastic steady state of  $\hat{X}_{t;\sigma}^{(1)}$  and  $\hat{X}_{t;\sigma}^{(2)}$  are both equal to zero, which ensures that the stochastic steady state of  $\hat{x}_{t;\sigma}^{(2)}$  is equal to the stochastic steady state of the original second-order perturbation approximation. In contrast, according to the formulation used in the literature the three variables,  $X_{t;\sigma}^{(2)}$ ,  $\tilde{X}_{t;\sigma}^{(1)}$ , and  $\tilde{X}_{t;\sigma}^{(2)}$  have three different steady state values. For the discussion in this section, this difference does not matter, because here we consider the case that  $\sigma \rightarrow 0$  and the three stochastic steady states would then converge to the non-stochastic steady state.<sup>45</sup>

The parameter  $\sigma$  plays two roles in the approximation. First, it controls the volatility of the driving process. Second, it affects the coefficients of the approximation. Here it is convenient to separate these two roles. We let  $\Omega$  indicate the value of volatility used to determine the coefficients of the perturbation approximation and we let  $\sigma$  indicate the value of the driving process. That is, we write the regular second-order perturbation

---

<sup>45</sup>As discussed in the main text, the motivation for our modification is the following. If higher-order perturbation introduces a correction for uncertainty, then it makes sense to apply this correction to *all* measures of the state variables that are introduced by the pruning procedure.

approximation as

$$X_{t;\Omega}^{(2)} = \bar{f}'_{x;\Omega} X_{t-1;\Omega}^{(2)} + \bar{f}'_{z;\Omega} \sigma \varepsilon_t + 0.5 \bar{f}''_{x^2;\Omega} \left( X_{t-1;\Omega}^{(2)} \right)^2 + 0.5 \bar{f}''_{z^2;\Omega} (\sigma \varepsilon_t)^2 + \bar{f}''_{xz;\Omega} X_{t-1;\Omega}^{(2)} \sigma \varepsilon_t \quad (65)$$

and the corresponding pruned perturbation approximation as

$$\hat{X}_{t;\Omega}^{(2)} = \bar{f}'_{x;\Omega} \hat{X}_{t-1;\Omega}^{(2)} + \bar{f}'_{z;\Omega} \sigma \varepsilon_t + 0.5 \bar{f}''_{x^2;\Omega} \left( \hat{X}_{t-1;\Omega}^{(1)} \right)^2 + 0.5 \bar{f}''_{z^2;\Omega} (\sigma \varepsilon_t)^2 + \bar{f}''_{xz;\Omega} \hat{X}_{t-1;\Omega}^{(1)} \sigma \varepsilon_t \quad (66a)$$

$$\hat{X}_{t;\Omega}^{(1)} = \bar{f}'_{x;\Omega} \hat{X}_{t-1;\Omega}^{(1)} + \bar{f}'_{z;\Omega} \sigma \varepsilon_t \quad (66b)$$

We know that the regular perturbation approximation given in Equation (65) is a second-order approximation if  $\sigma$  and  $\Omega$  approach zero.<sup>46</sup> Consequently, if the perturbation approximation given in Equation (66) approaches the perturbation approximation given in Equation (65) as  $\sigma \rightarrow 0$ , then it is a second-order approximation of the truth if  $\sigma$  and  $\Omega$  approach zero.

We will show that our pruned perturbation approximation converges to the regular perturbation approximation for fixed  $\Omega$ , which implies that our procedure correctly approximates the correction that higher-order perturbation introduces for uncertainty (indicated by  $\Omega$ ) at least as  $\sigma \rightarrow 0$ . For the regular pruning formulation this is only true if both  $\Omega \rightarrow 0$  and  $\sigma \rightarrow 0$ .

The difference between  $X_{t;\Omega}^{(2)}$  and  $\hat{X}_{t;\Omega}^{(2)}$  is equal to

$$\begin{aligned} X_{t;\Omega}^{(2)} - \hat{X}_{t;\Omega}^{(2)} &= \bar{f}'_{x;\Omega} \left( X_{t-1;\Omega}^{(2)} - \hat{X}_{t-1;\Omega}^{(2)} \right) \\ &+ 0.5 \bar{f}''_{x^2;\Omega} \left( \left( X_{t-1;\Omega}^{(2)} \right)^2 - \left( \hat{X}_{t-1;\Omega}^{(1)} \right)^2 \right) \\ &+ \bar{f}''_{xz;\Omega} \sigma \varepsilon_t \left( \sigma X_{t-1;\Omega}^{(2)} - \sigma \hat{X}_{t-1;\Omega}^{(1)} \right) \end{aligned} \quad (67)$$

To see that  $X_{t;\Omega}^{(2)} - \hat{X}_{t;\Omega}^{(2)} = O(\sigma^3)$ , first note that

$$X_{t-1;\Omega}^{(2)} - \hat{X}_{t-1;\Omega}^{(1)} = O(\sigma^2), \quad (68)$$

which means that

$$\sigma X_{t-1;\Omega}^{(2)} - \sigma \hat{X}_{t-1;\Omega}^{(1)} = O(\sigma^3). \quad (69)$$

---

<sup>46</sup>As in Appendix B.1, one needs to assume that the initial condition converges to the steady state at the appropriate rate as  $\sigma$  approaches zero.

Moreover, since  $X_{t-1;\Omega}^{(1)}$  is  $O(\sigma^2)$  and  $X_{t-1;\Omega}^{(2)}$  is  $O(\sigma^3)$ , it is also true that

$$\left(X_{t-1;\Omega}^{(2)}\right)^2 - \left(\hat{X}_{t-1;\Omega}^{(1)}\right)^2 = O(\sigma^3). \quad (70)$$

Consequently, the only term remaining is  $X_{t-1;\Omega}^{(2)} - \hat{X}_{t-1;\Omega}^{(2)}$ . But given the last two results this will be  $O(\sigma^3)$  as long as the difference in the initial conditions is  $O(\sigma^3)$  as in Equation (58).

### B.3.2 Convergence of higher-order pruning

The discussion above easily extends to the case for higher-order pruning. But one should be careful in specifying the formulation for higher-order pruned perturbation. We make this clear using a simple example.

Let the true law of motion be given by

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-1}^2 + \rho_3 x_{t-1}^3 + \sigma \varepsilon_t, \quad (71a)$$

$$\mathbb{E}[\varepsilon_t^2] = 1. \quad (71b)$$

As discussed in Section 4.3, the third-order pruned perturbation solution is given by

$$\hat{x}_t^{(3)} = \rho_1 \hat{x}_{t-1}^{(3)} + \rho_2 \left(\hat{x}_{t-1}^{(2)}\right)^2 + \rho_3 \left(\hat{x}_{t-1}^{(1)}\right)^3 + \sigma \varepsilon_t \quad (72)$$

$$\hat{x}_t^{(2)} = \rho_1 \hat{x}_{t-1}^{(2)} + \rho_2 \left(\hat{x}_{t-1}^{(1)}\right)^2 + \sigma \varepsilon_t \quad (73)$$

$$\hat{x}_t^{(1)} = \rho_1 \hat{x}_{t-1}^{(1)} + \sigma \varepsilon_t \quad (74)$$

Using the same logic as used in Section B.3.1 for second-order pruning, it is easy to show that

$$x_t - \hat{x}_t^{(3)} = O(\sigma^4). \quad (75)$$

Now consider the following alternative formulation:

$$\tilde{x}_t^{(3)} = \rho_1 \tilde{x}_{t-1}^{(3)} + \rho_2 \left(\tilde{x}_{t-1}^{(1)}\right)^2 + \rho_3 \left(\tilde{x}_{t-1}^{(1)}\right)^3 + \sigma \varepsilon_t \quad (76a)$$

$$\tilde{x}_t^{(1)} = \rho_1 \tilde{x}_{t-1}^{(1)} + \sigma \varepsilon_t \quad (76b)$$

This formulation does not generate a proper third-order approximation. To see that

$$x_t - \tilde{x}_t^{(3)} \neq O(\sigma^4) \quad (77)$$

consider the following example. Suppose that

$$x_0 = 0, \quad (78)$$

$$\varepsilon_t = 1 \text{ if } t = 1, \quad (79)$$

$$\varepsilon_t = 0 \text{ if } t > 1. \quad (80)$$

Table 3 reports the time paths for  $x_t$  and  $\tilde{x}_t^{(3)}$ . The term in bold,  $2\rho_1\rho_2\sigma^3$ , in the third period makes clear that the difference between  $x_t$  and  $\tilde{x}_t^{(3)}$  is not of order  $O(\sigma^4)$ .

## C Implementation of the perturbation-plus procedure

In this appendix, we describe a faster version of the perturbation-plus procedure and we describe in detail how we implemented the perturbation-plus procedure.

### C.1 Faster implementation

An important factor that slows down the perturbation procedure is that there is no analytical solution to Equation (30), which we repeat here for convenience.

$$0 = \tilde{\mathbb{E}} \left[ H(\tilde{f}_{1\text{st}}(x, z_{+1}), x, x_{-1}, z_{+1}, z) \right]. \quad (81)$$

The objective of the alternative formulation is to avoid using a nonlinear equation solver. Typically,  $x$  shows up more than once in  $H(\cdot)$  and typically it is possible to find an analytical expression for  $x$  in terms of the other variables and  $x$  itself. If that is the case, then we can rewrite Equation (81) as<sup>47</sup>

$$x = \tilde{\mathbb{E}} \left[ G(\tilde{f}_{1\text{st}}(x, z_{+1}), x, x_{-1}, z_{+1}, z) \right]. \quad (82)$$

Since this is still an equation in  $x$ , we have not made any progress. The idea is to use  $\tilde{f}_{1\text{st}}(\cdot)$  not only for  $x_{+1}$  but also for the value of  $x$  inside  $G(\cdot)$ .<sup>48</sup> The one-step ahead

---

<sup>47</sup>See Section C.3 for an example.

<sup>48</sup>The difference between the original perturbation-plus procedure and its modification is very similar to the difference between time iteration and fixed-point iteration. See Chapter 17 in Judd (1998) for a discussion on these iteration schemes.

modification is then defined as

$$x = \widehat{f}_{+1}(x_{-1}, z) = \widetilde{\text{E}}_t \left[ G(\widetilde{f}_{1^{\text{st}}}(x_{-1}, z), z_{+1}), \widetilde{f}_{1^{\text{st}}}(x_{-1}, z), x_{-1}, z_{+1}, z \right]. \quad (83)$$

The two-step ahead modification is defined as

$$x = \widehat{f}_{+2}(x_{-1}, z) = \widetilde{\text{E}}_t \left[ G(\widehat{f}_{+1}(\widehat{f}_{+1}(x_{-1}, z), z_{+1}), \widehat{f}_{+1}(x_{-1}, z), x_{-1}, z_{+1}, z \right]. \quad (84)$$

Iteration on this scheme leads to the  $J$ -step ahead modification of the first-order perturbation solution.

## C.2 Perturbation-plus and the neo-classical growth model

The first-order conditions for the neo-classical growth model are given by

$$c_t + k_t = e^{z_t} k_{t-1}^\alpha + (1 - \delta)k_{t-1} \quad (85)$$

and

$$1 = \text{E}_t \left[ \left( \frac{c_{t+1}}{c_t} \right)^{-\gamma} (\alpha e^{z_{t+1}} k_t^{\alpha-1} + 1 - \delta) \right]. \quad (86)$$

The objective is to solve for  $k_t$  given values of  $k_{t-1}$  and  $z_t$ . We denote the solution as  $k_t = \widehat{f}_{+1}(k_{t-1}, z_t)$ . The solution for  $k_t$  and  $c_t$  are obtained using the following two equations

$$c_t + k_t = e^{z_t} k_{t-1}^\alpha + (1 - \delta)k_{t-1} \quad (87a)$$

$$1 = \widetilde{\text{E}}_t \left[ \left( \frac{e^{z_{t+1}} k_t^\alpha + (1 - \delta)k_t - \widetilde{f}_{1^{\text{st}}}(k_t, z_{t+1})}{c_t} \right)^{-\gamma} (\alpha e^{z_{t+1}} k_t^{\alpha-1} + 1 - \delta) \right]. \quad (87b)$$

where  $\widetilde{f}_{1^{\text{st}}}(k_t, z_{t+1})$  is the first-order perturbation solution for  $k_{t+1}$ . The conditional expectation  $\widetilde{\text{E}}_t[\cdot]$  is the numerical approximation to  $\text{E}_t[\cdot]$  using Gaussian-Hermite quadrature with five quadrature nodes.

The procedure described in the last paragraph is the one-step ahead perturbation-plus procedure, because the behavior of next period's variables is described using the first-order perturbation procedure. The two-step ahead perturbation-plus approximation,  $\widehat{f}_{+2}(\cdot)$ , is the solution to Equation (87) with  $\widetilde{f}_{1^{\text{st}}}(\cdot)$  replaced by  $\widehat{f}_{+1}(\cdot)$ .

### C.3 Perturbation-plus and the matching model

The objective is to solve for  $\lambda_t$  given the values of the state variables,  $n_{t-1}$  and  $z_t$ . We denote this solution by

$$\lambda_t = \widehat{f}_{+1}(n_{t-1}, z_t). \quad (88)$$

We solve for  $\lambda_t$  from a system of five equations in five unknowns. The endogenous variables are  $\lambda_t$ ,  $p_{f,t}$ ,  $v_t$ ,  $c_t$ , and  $n_t$ . The five equations are

$$\lambda_t = \beta E \left[ \left( \frac{\widetilde{c}_{t+1}(n_t, z_{t+1})}{c_t} \right)^{-\gamma} \left( \alpha e^{z_{t+1}} n_t^{\alpha-1} - w + (1 - \rho_n) \widetilde{f}_{1st}(n_t, z_{t+1}) \right) \right], \quad (89)$$

and Equations (7), (10), (12), and (14). For this to be a system in five unknowns, we have to take a stand on how to determine  $\widetilde{c}_{t+1}(n_t, z_{t+1})$ . Since this is *next period's* consumption, we could use the first-order perturbation approximation. Instead we use for  $\widetilde{c}_{t+1}(n_t, z_{t+1})$  the value that is implicitly defined by Equations (7), (10), (12), and (14) for  $t + 1$  and

$$\lambda_{t+1} = \widetilde{f}_{1st}(n_t, z_{t+1}). \quad (90)$$

Thus, we only use the first-order perturbation solution for  $\lambda_{t+1}$  and all the other variables are obtained using the true model equations.

The two-step ahead perturbation-plus approximation is defined analogously with  $\widetilde{f}_{1st}(\cdot)$  replaced by  $\widehat{f}_{+1}(\cdot)$ .

**Simplification.** The functions  $\widehat{f}_{+J}(\cdot)$  are only implicitly defined and its values are calculated using a nonlinear equation solver. Consequently, the cost of the algorithm increases sharply with  $J$ . In Appendix C.1, we discussed a simplification that reduces the costs substantially. Here we discuss how this is implemented.

Instead of solving for  $\lambda_t$ ,  $n_t$  and  $c_t$  simultaneously, we use the following procedure. First, calculate "temporary" values  $n_t$  and  $c_t$  using Equations (7), (10), (12), (14), and

$$\lambda_t = \widetilde{f}_{1st}(n_{t-1}, z_t). \quad (91)$$

Denote the solutions for  $n_t$  and  $c_t$  as  $n_t^{\text{temp}}$  and  $c_t^{\text{temp}}$ . As above, the value for  $\lambda_t$  is

calculated from

$$\lambda_t = \beta \mathbb{E} \left[ \left( \frac{\tilde{c}_{t+1}(n_t^{\text{temp}}, z_{t+1})}{c_t^{\text{temp}}} \right)^{-\gamma} \left( \begin{array}{c} \alpha e^{z_{t+1}} (n_t^{\text{temp}})^{\alpha-1} - w \\ + (1 - \rho_n) \tilde{f}_{1\text{st}}(n_t^{\text{temp}}, z_{t+1}) \end{array} \right) \right]. \quad (92)$$

With  $\tilde{c}_{t+1}(n_t^{\text{temp}}, z_{t+1})$  defined as above, we have an analytical expression for  $\lambda_t$ . The  $n_t^{\text{temp}}$  variable is only used to calculate  $\lambda_t$ . Given the solution for  $\lambda_t$ , the actual value for  $n_t$  is then obtained from the Equations (7), (10), (12), and (14), without making any further approximation. To calculate  $\hat{f}_{+J}(\cdot)$  one would use  $\hat{f}_{+J-1}(\cdot)$  instead of  $\tilde{f}_{1\text{st}}(\cdot)$  both to calculate  $n_t^{\text{temp}}$ ,  $c_t^{\text{temp}}$ , and to calculate  $\lambda_{t+1}$  in the Euler equation.

#### C.4 Perturbation-plus and the modified Deaton model

In each period of the simulation, i.e., given the value for cash on hand,  $x_t$ , we use an equation solver to calculate  $a_t$  from the Euler equation. We use Gaussian-Hermite quadrature to calculate the conditional expectation on the right-hand side of the Euler equation.<sup>49</sup>

For the one-step ahead modification, we use

$$c_{t+1} = a_t + e^{z_{t+1}} - \frac{a_{t+1}}{1+r} \quad (93)$$

$$\approx a_t + e^{z_{t+1}} - \frac{\tilde{f}_{1\text{st}}(a_t + e^{z_{t+1}})}{1+r} \quad (94)$$

to calculate the realizations for consumption. This procedure defines the function  $a_t = \hat{f}_{+1}(x_t)$ .

For the two-step ahead modification, we use  $a_{t+1} = \hat{f}_{+1}(x_{t+1})$  in the expression for consumption above. For each Gaussian-Hermite node, i.e., for each potential value of  $a_t + e^{z_{t+1}}$ , we use an equation solver to calculate  $a_{t+1}$  from tomorrow's Euler equation and on the right-hand side we use  $a_{t+2} = \tilde{f}_{1\text{st}}(x_{t+2})$ . Since  $\tilde{f}_{1\text{st}}(\cdot)$  is only implicitly defined, we have to use an equation solver to solve for  $a_{t+1}$  for each quadrature node for  $z_{t+1}$ .

Although computationally expensive, it is easy to iterate on this procedure to calculate the  $J$ -step ahead modification.

---

<sup>49</sup>For the number of quadrature nodes we considered values between five and thirty and found that the results were robust to changing this number.



## D Accuracy of our projection solutions

In this appendix, we document that the projection solutions that serve as a stand-in for the truth are very accurate.

### D.1 Projection solution for matching model

We obtained a very accurate solution for the matching model using the following algorithm based on projection methods. We parameterized the policy function for the Lagrange multiplier  $\lambda$  by a linear spline that satisfies the Euler equation on each grid point. We have used 10,000 equidistant grid points for  $n_{-1}$  ranging from 0.6 to 0.99. The other state variable,  $z$ , can take on two values, namely  $-\zeta$  and  $+\zeta$ .

We used fixed-point iteration and the algorithm does the following at the  $i^{\text{th}}$  iteration. Starting point of the  $i^{\text{th}}$  iteration is the policy function from the last iteration, namely  $\lambda = f^{(i-1)}(n_{-1}, z)$ . Given this policy function it is straightforward to solve for the other variables. At grid point  $j$ , i.e., for given values of  $n_{-1,(j)}$  and  $z_{(j)}$ , the value for  $\lambda$  is given by

$$\lambda_{(j)} = \beta \mathbb{E} \left[ \left( \frac{c(n_{(j)}, z_{+1})}{c_{(j)}} \right)^{-\gamma} \left( \alpha e^{z_{+1}} n_{(j)}^{\alpha-1} - w + (1 - \rho_n) \lambda(n_{(j)}, z_{+1}) \right) \right]. \quad (95)$$

Integrating over the possible realizations for  $z_{+1}$  is trivial given that  $z$  has discrete support. If  $\lambda(n_{(j)}, z_{+1})$  and  $c(n_{(j)}, z_{+1})$  are determined using  $f^{(i-1)}(n_{(j)}, z_{+1})$ , then one can solve for  $\lambda_{(j)}$ ,  $n_{(j)}$ ,  $v_{(j)}$ ,  $p_{f,(j)}$ , and  $c_{(j)}$  by combining Equation (95) with Equations (7), (10), (12), and (14). This would be time iteration. To simplify the algorithm we use fixed-point iteration. The benefit of time iteration is that it has better convergence properties, but with the appropriate choice of the dampening parameter, the algorithm also converged with fixed-point iteration. We implemented fixed-point iteration as follows. First, calculate  $\lambda_{(j)}^{\text{temp}} = f^{(i-1)}(n_{-1,(j)}, z_{(j)})$ . Use this value to calculate  $n_{(j)}^{\text{temp}}$ . Next, solve for  $\lambda$  using

$$\lambda_{(j)} = \beta \mathbb{E} \left[ \left( \frac{c(n_{(j)}^{\text{temp}}, z_{+1})}{c_{(j)}} \right)^{-\gamma} \left( \alpha e^{z_{+1}} (n_{(j)}^{\text{temp}})^{\alpha-1} - w + (1 - \rho_n) \lambda(n_{(j)}^{\text{temp}}, z_{+1}) \right) \right], \quad (96)$$

where the values for  $c_{+1}$  and  $\lambda_{+1}$  are based on  $f^{(i-1)}(\cdot)$ . In principle one could set

$f^{(i)}(n_{-1,(j)}, z_{(j)})$  equal to  $\lambda_{(j)}$ . But convergence may require a dampening factor, that is, to take a weighted average between  $\lambda_{(j)}$  and  $f^{(i-1)}(n_{-1,(j)}, z_{(j)})$ . We iterate on this scheme until the maximum absolute difference is less than 1E-12.

As documented in Table 1, the errors made in the dynamic Euler equation accuracy test are minuscule.

## D.2 Projection solution for modified Deaton model

We obtained an accurate solution for the modified Deaton model using projection techniques. The details are as follows. We parameterized the asset policy function by a linear spline that satisfies the Euler equation on each grid point. We use time iteration and we use the endogenous grid points algorithm of Carroll (2006). The advantage of time iteration (compared to fixed-point iteration) is that it has better convergence properties. The advantage of endogenous grid points is that there is an analytical solution for the variables when using time iteration. The disadvantage of using time iteration and endogenous grid points is that we have to specify a grid for both  $a_t$  and  $z_t$ , while strictly speaking there is only one state variable, namely cash on hand,  $a_{t-1} + e^{z_t}$ .

We used 1001 equidistant grid points for  $a_t$  ranging from  $-0.2$  to  $2$  and we used 1001 equidistant grid points for  $z_t$  ranging from  $0$  to  $3$ .

The projection algorithm is based on an iterative scheme that does the following at the  $i^{\text{th}}$  iteration. Starting point at the  $i^{\text{th}}$  iteration is the policy function from the last iteration, namely  $a = f^{(i-1)}(a_{-1} + e^z)$ . At grid point  $j$ , i.e., for given values of  $a_{(j)}$  and  $z_{(j)}$ , the value for  $a_{-1}$  is solved from the Euler equation. The conditional expectation is approximated using Gaussian-Hermite quadrature with 30 nodes. The possible realizations for  $c_{+1}$  are given by

$$c_{+1} = a_{(j)} + e^{z_{+1}} - \frac{f^{(i-1)}(a_{(j)} + e^{z_{+1}})}{1 + r}. \quad (97)$$

This leads to a set of combinations of  $a_{-1}$ ,  $a$ , and  $z$  which gives  $a = f(a_{-1} + e^z)$ . We iterate on this scheme until the maximum absolute difference between the values of  $f^{(i-1)}$  and  $f^{(i)}$  is less than 1E-7.

As documented in Table 1, the errors made in the dynamic Euler equation accuracy

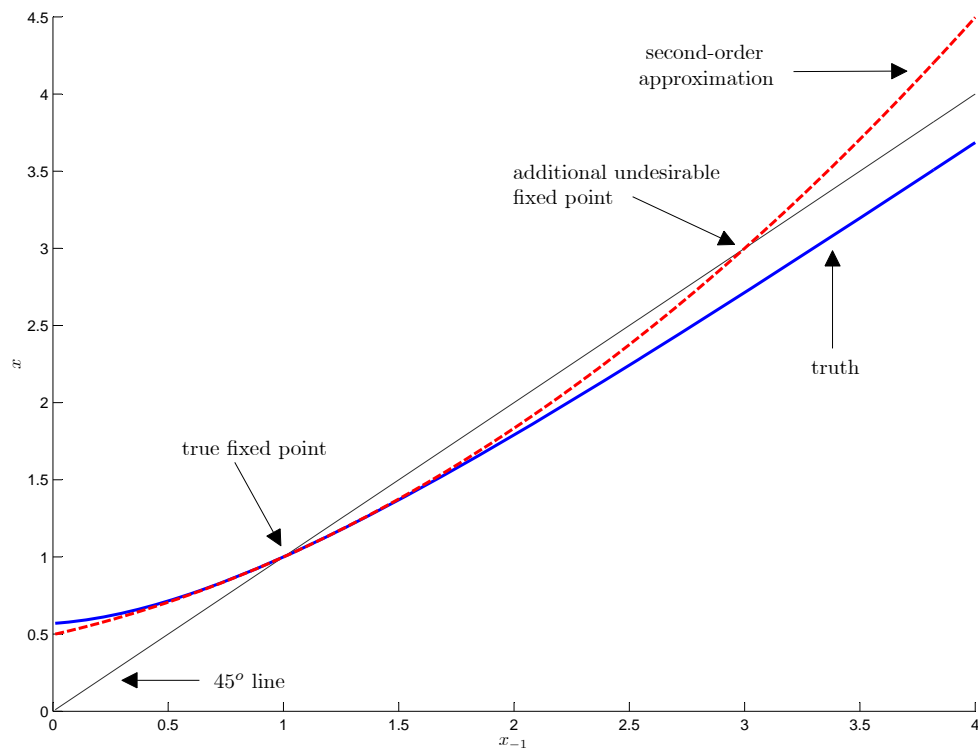
test are not quite as small as those for the matching model. This model is more difficult to solve, given that the variance is much higher. Nevertheless, the results are still good. The maximum error is 0.1% and the average error is 0.008%. These errors are minuscule relative to the errors of the perturbation based methods.

## References

- ANDREASEN, M. M. (2008): “Non-Linear DSGE Models, The Central Difference Kalman Filter, and The Mean Shifted Particle Filter,” Unpublished manuscript, Aarhus.
- CARROLL, C. D. (2006): “The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems,” *Economic Letters*, 91, 312–320.
- DE WIND, J. (2008): “Punishment Functions,” Unpublished manuscript, University of Amsterdam.
- DEATON, A. (1991): “Saving and Liquidity Constraints,” *Econometrica*, 59, 1221–1248.
- DEN HAAN, W. J., AND J. DE WIND (2009): “How Well-Behaved are Higher-Order Perturbation Solutions?,” Unpublished manuscript, University of Amsterdam.
- DEN HAAN, W. J., AND G. KALTENBRUNNER (2009): “Anticipated Growth and Business Cycles in Matching Models,” *Journal of Monetary Economics*, 56, 309–327.
- DEN HAAN, W. J., AND P. RENDAHL (2010): “Solving the Incomplete Markets Model with Aggregate Uncertainty Using Explicit Aggregation,” *Journal of Economic Dynamics and Control*, 34, 69–78.
- DOH, T. (2009): “Yield Curve in an Estimated Nonlinear Macro Model,” Federal Reserve Bank of Kansas City Working Paper RWP 09-04.
- FAHR, S., AND F. SMETS (2008): “Downward Wage Rigidities and Optimal Monetary Policy in a Monetary Union,” Unpublished working paper, European Central Bank.

- FAIR, R. C., AND J. B. TAYLOR (1983): “Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models,” *Econometrica*, 51, 1169–1185.
- FERNÁNDEZ-VILLAYERDE, J., P. GUERRÓN-QUINTANA, J. F. RUBIO-RAMÍREZ, AND M. URIBE (2009): “Risk Matters: The Real Effects of Volatility Shock,” Unpublished manuscript, University of Pennsylvania.
- GAGNON, J. E. (1990): “Solving the Stochastic Growth Model by Deterministic Extended Path,” *Journal of Business and Economic Statistics*, 8, 35–36.
- HAGEDORN, M., AND I. MANOVSKII (2008): “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited,” *American Economic Review*, 98, 1692–1706.
- JUDD, K. L. (1998): *Numerical Methods in Economics*. The MIT Press, Cambridge, Massachusetts.
- KIM, J., S. KIM, E. SCHAUMBURG, AND C. A. SIMS (2008): “Calculating and Using Second-Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models,” *Journal of Economic Dynamics and Control*, 32, 3397–3414.
- KRUSELL, P., AND A. A. SMITH, JR. (1998): “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, 106, 867–896.
- LOMBARDO, G. (2010): “Approximating DSGE Models by Series Expansions,” Unpublished manuscript, European Central Bank.
- PETRONGOLO, B., AND C. PISSARIDES (2001): “Looking Into the Black Box: A Survey of the Matching Function,” *Journal of Economic Literature*, 39, 390–431.

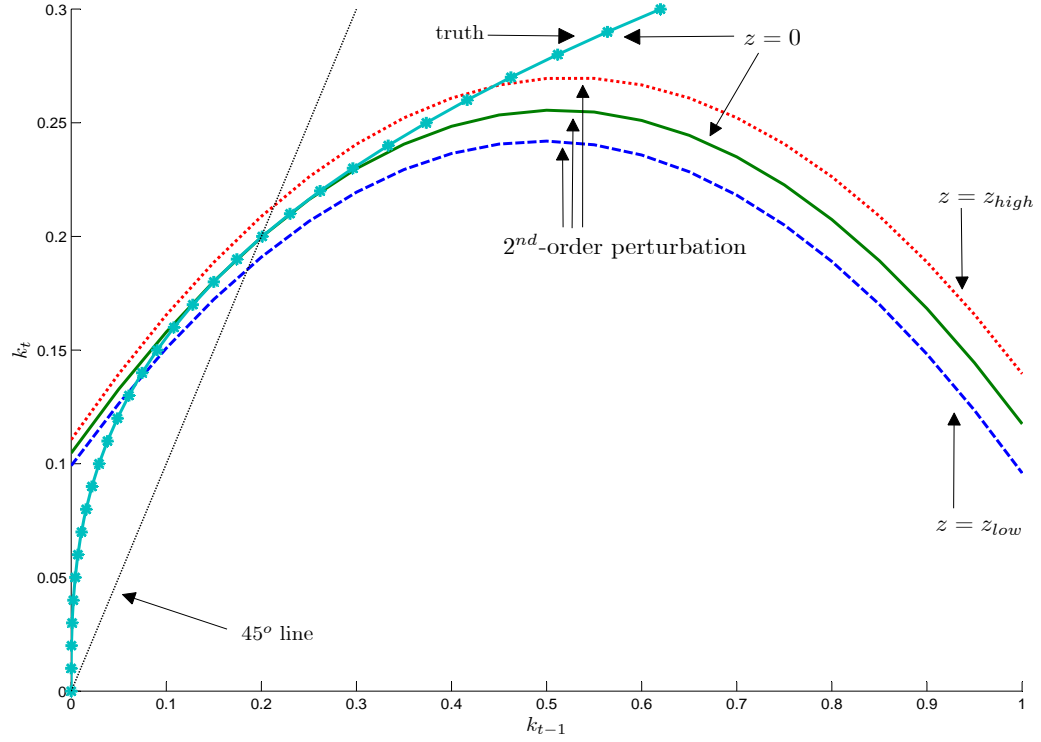
Figure 1: Perturbation approximations and instability



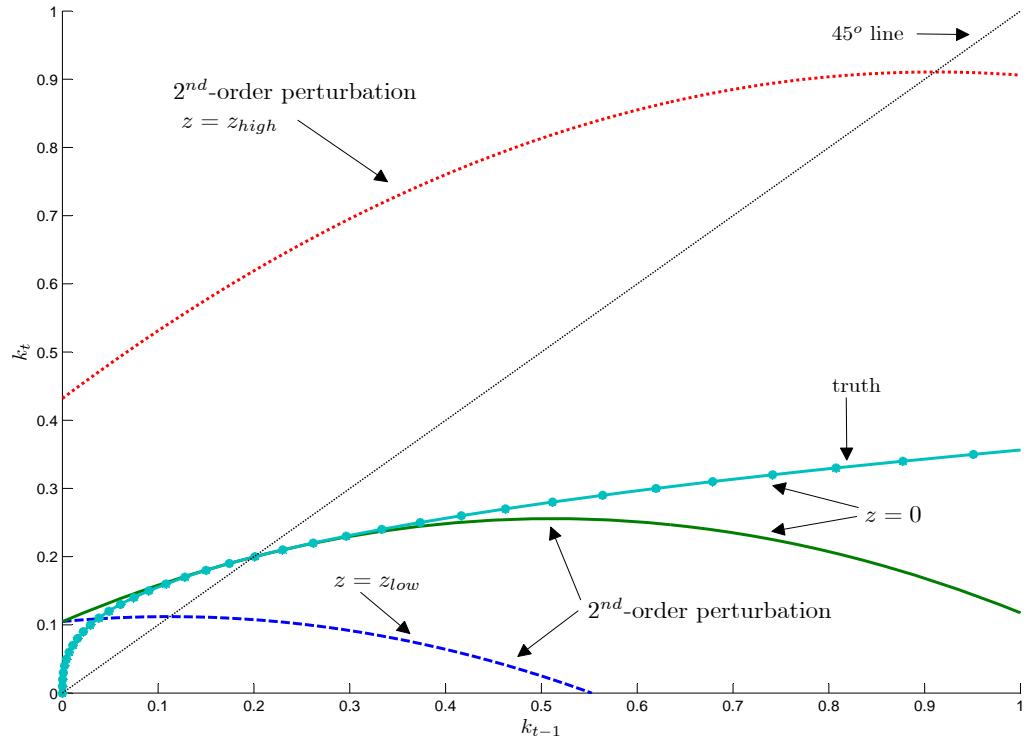
Notes: This figure plots the function  $f(x_{-1})$  described in Section 1 and its second-order Taylor-series approximation.

Figure 2: 2<sup>nd</sup>-order perturbation approximations for Brock-Mirman model

A: Approximation for  $k(k_{-1}, z)$ ;  $\sigma_z = 0.007$

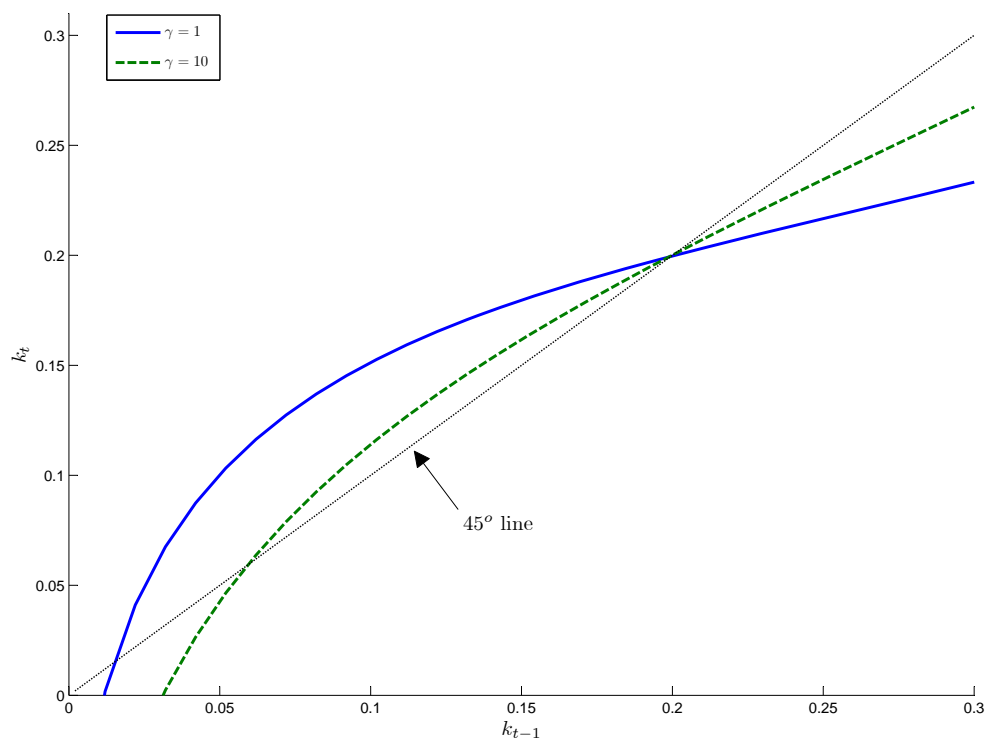


B: Approximation for  $k(k_{-1}, z)$ ;  $\sigma_z = 0.2$



Notes: This figure plots the 2<sup>nd</sup>-order perturbation approximations for the Brock-Mirman model (in levels not in logarithms) for different values of  $z$ .

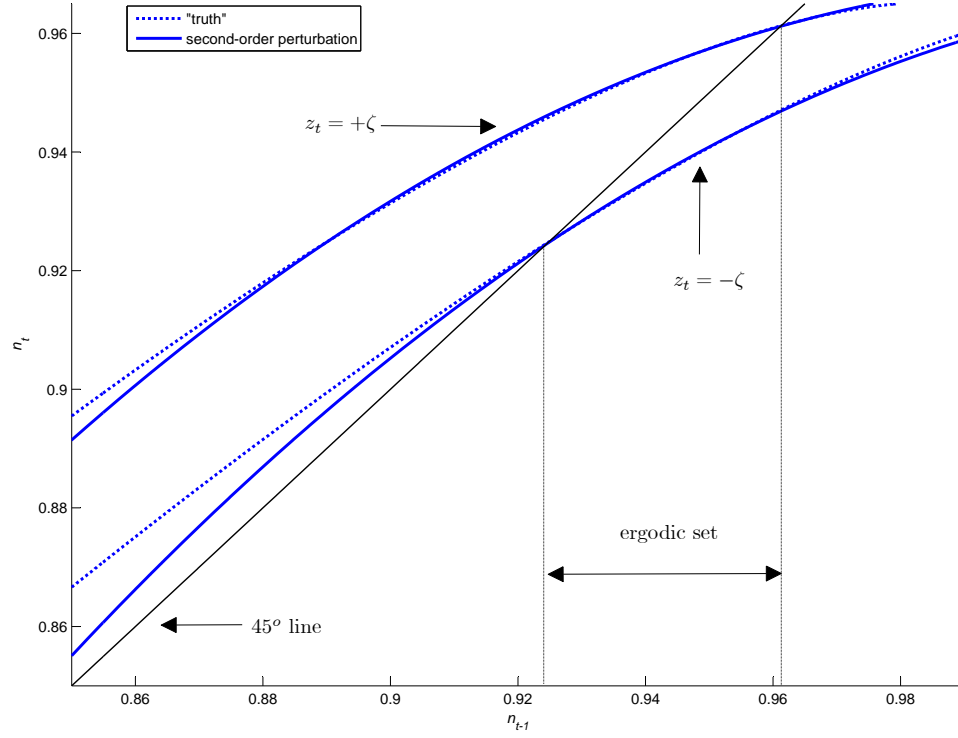
Figure 3: 2<sup>nd</sup>-order perturbation approximations for neo-classical growth model; consumption perturbation solution and capital solved from true budget constraint



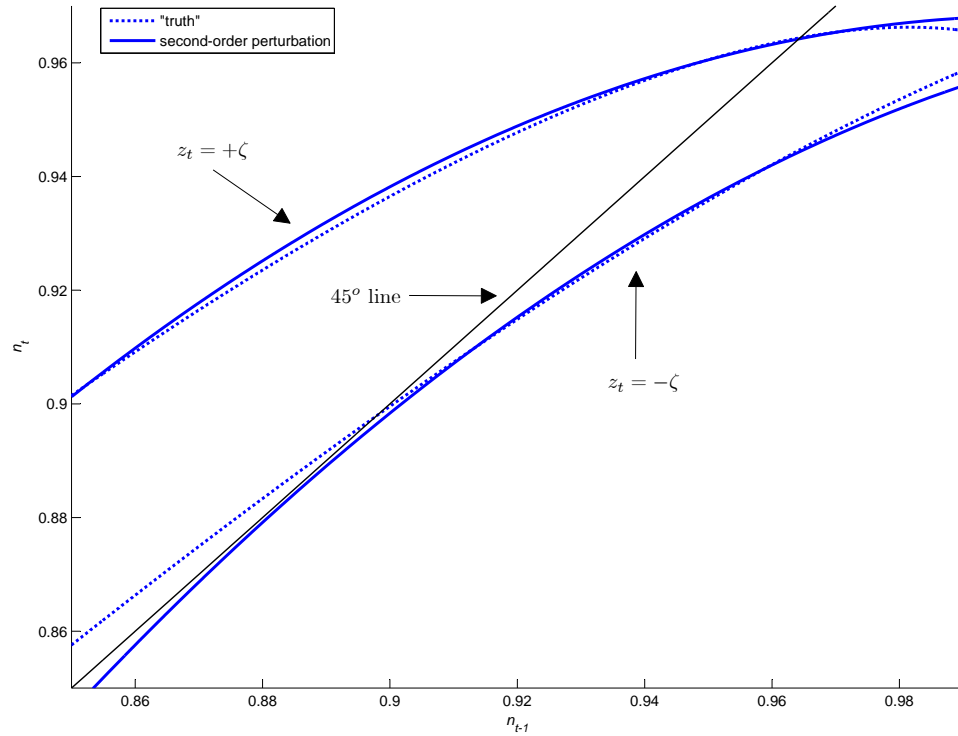
Notes: When  $\gamma = 1$ , then the model is the Brock-Mirman model.

Figure 4: "Truth" and 2<sup>nd</sup>-order perturbation approximation for matching model

A:  $w = 0.96$ .



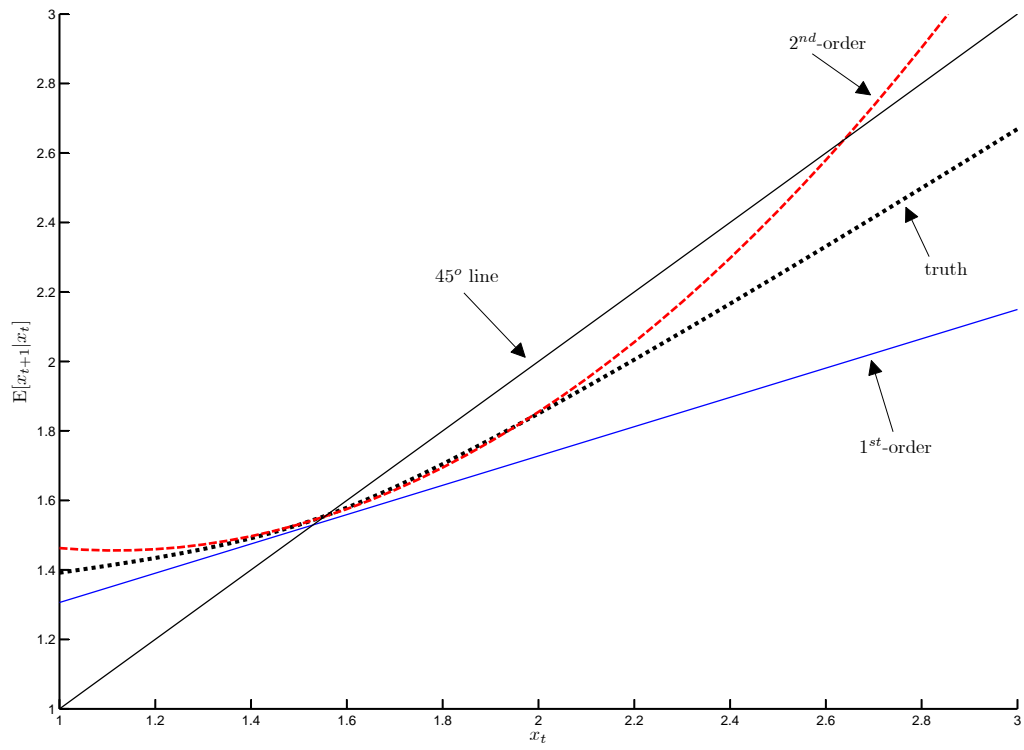
B:  $w = 0.973$ .



Notes: The point of this figure is to show that the second-order perturbation solution for  $z = -\zeta$  does not have a fixed point unless  $w$  is sufficiently low.

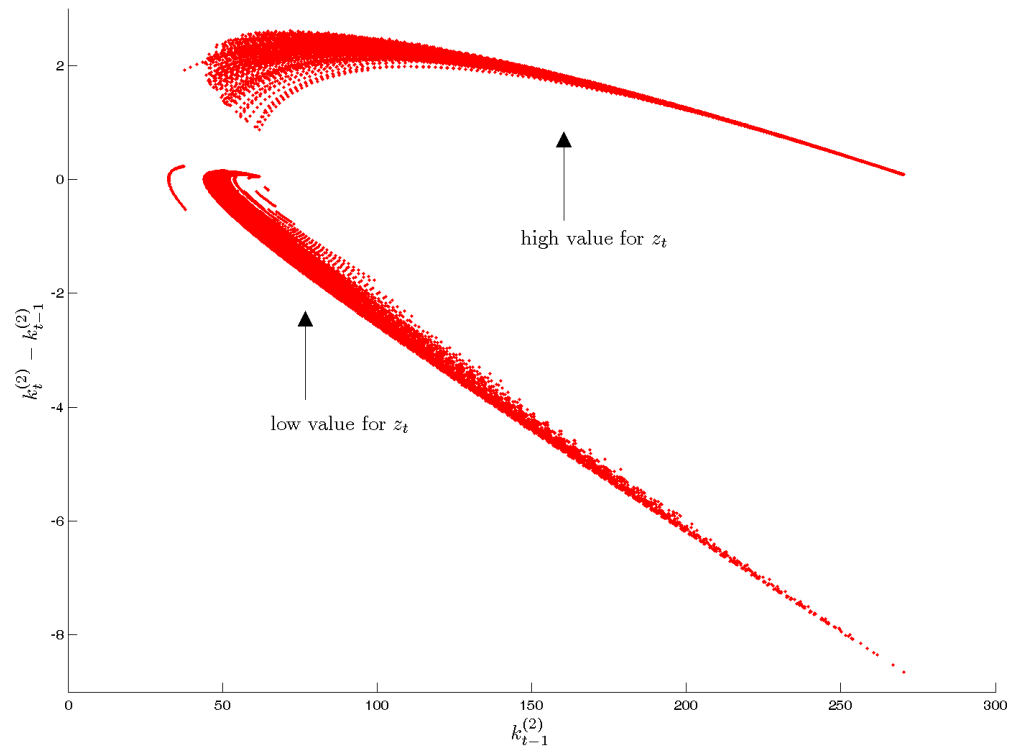


Figure 5: 2<sup>nd</sup>-order perturbation approximations for the Deaton model



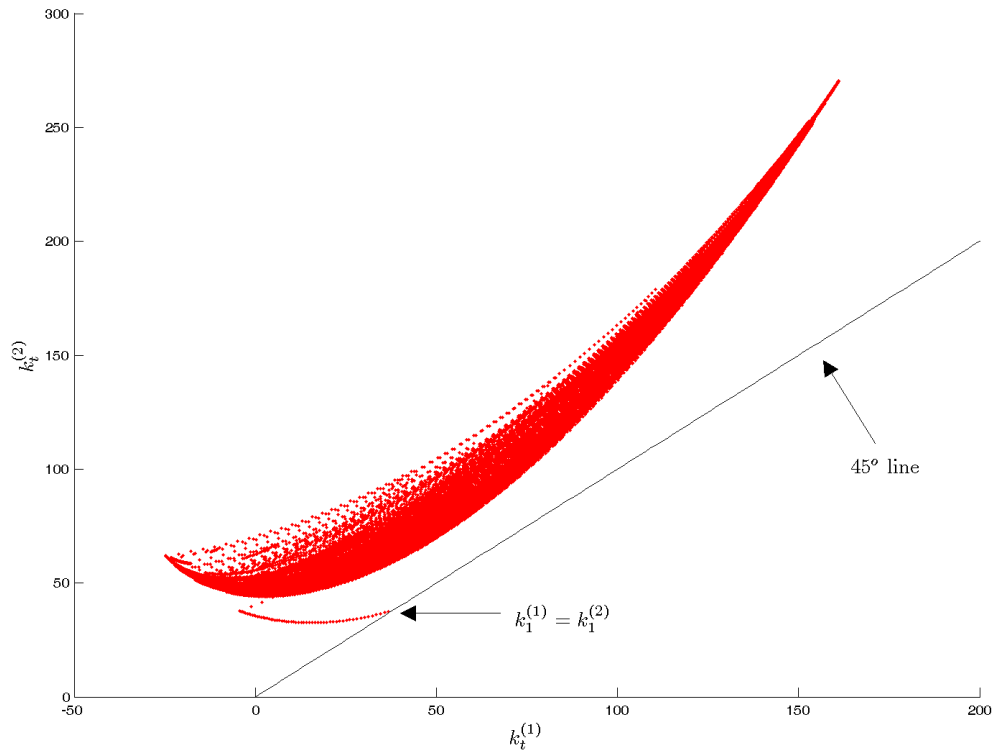
Notes: The point of this figure is to show that the second-order perturbation approximation for the policy function of the Deaton model is not globally stable.

Figure 6: Data simulated with 2<sup>nd</sup>-order pruned perturbation approximation;  
 $k_t^{(2)} - k_{t-1}^{(2)}$  as a "function" of  $k_{t-1}^{(2)}$



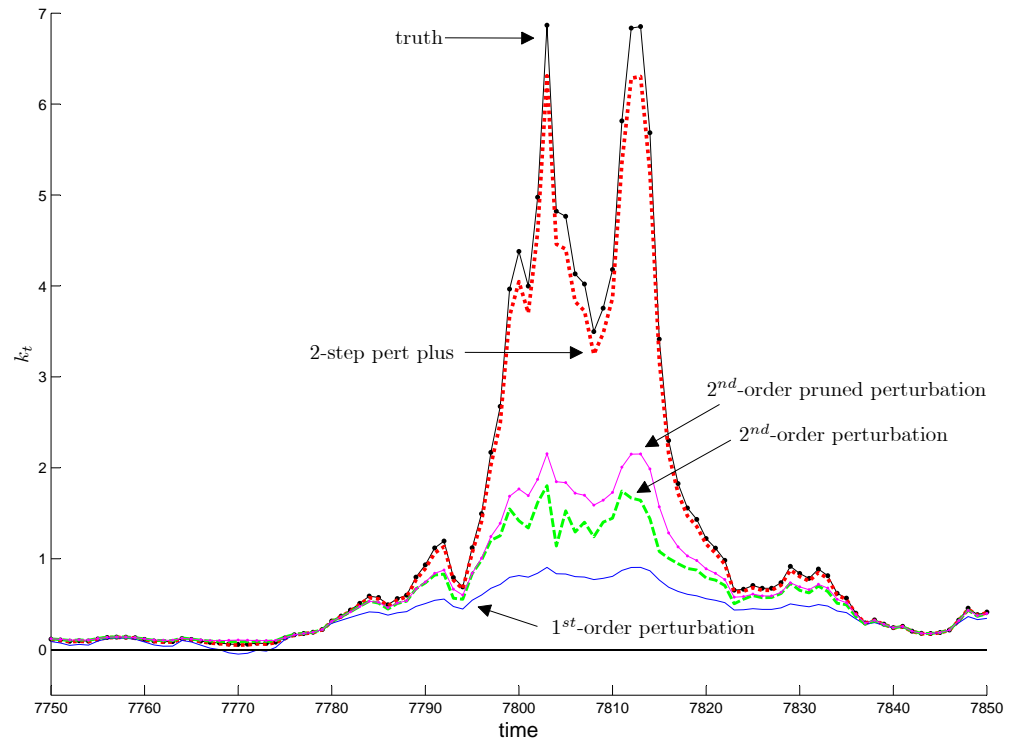
Notes:  $k_t^{(2)}$  is the capital stock generated by the pruning procedure.  $k_t^{(2)}$  is not a function of  $k_{t-1}^{(2)}$  and  $z_t$ , since the pruning procedure introduces additional state variables.

Figure 7: Data simulated with 2<sup>nd</sup>-order pruned perturbation approximation;  
 $k_t^{(2)}$  versus  $k_t^{(1)}$



Notes:  $k_t^{(2)}$  is the capital stock generated by the pruning procedure and  $k_t^{(1)}$  is the value of the auxiliary state variable also generated by the pruning procedure.

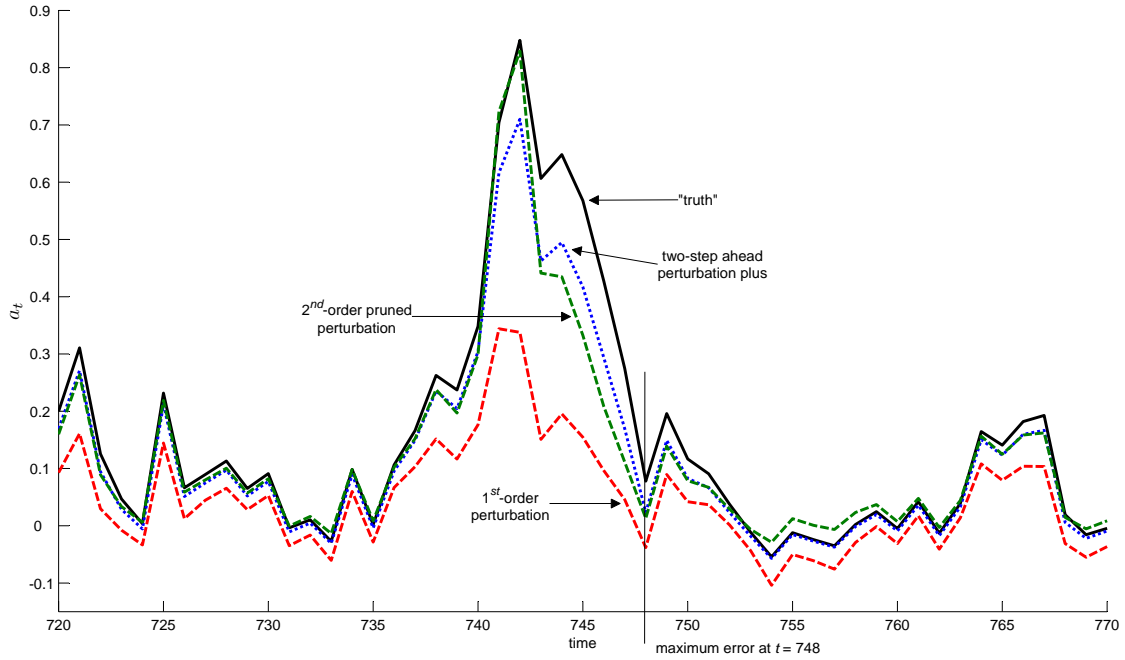
Figure 8: Time paths for the Brock-Mirman model



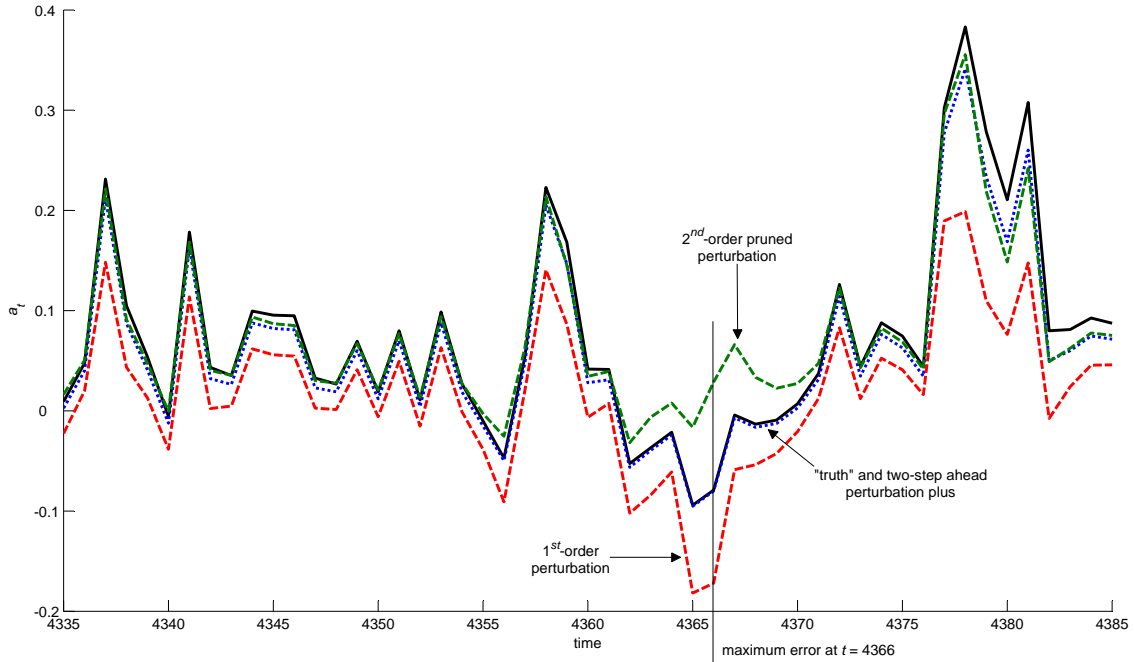
Notes: This graphs plots the different approximations when the 2-step ahead perturbation-plus approximation attains the largest error.

Figure 9: Time paths for the modified Deaton model

A: Maximum error for two-step ahead perturbation plus

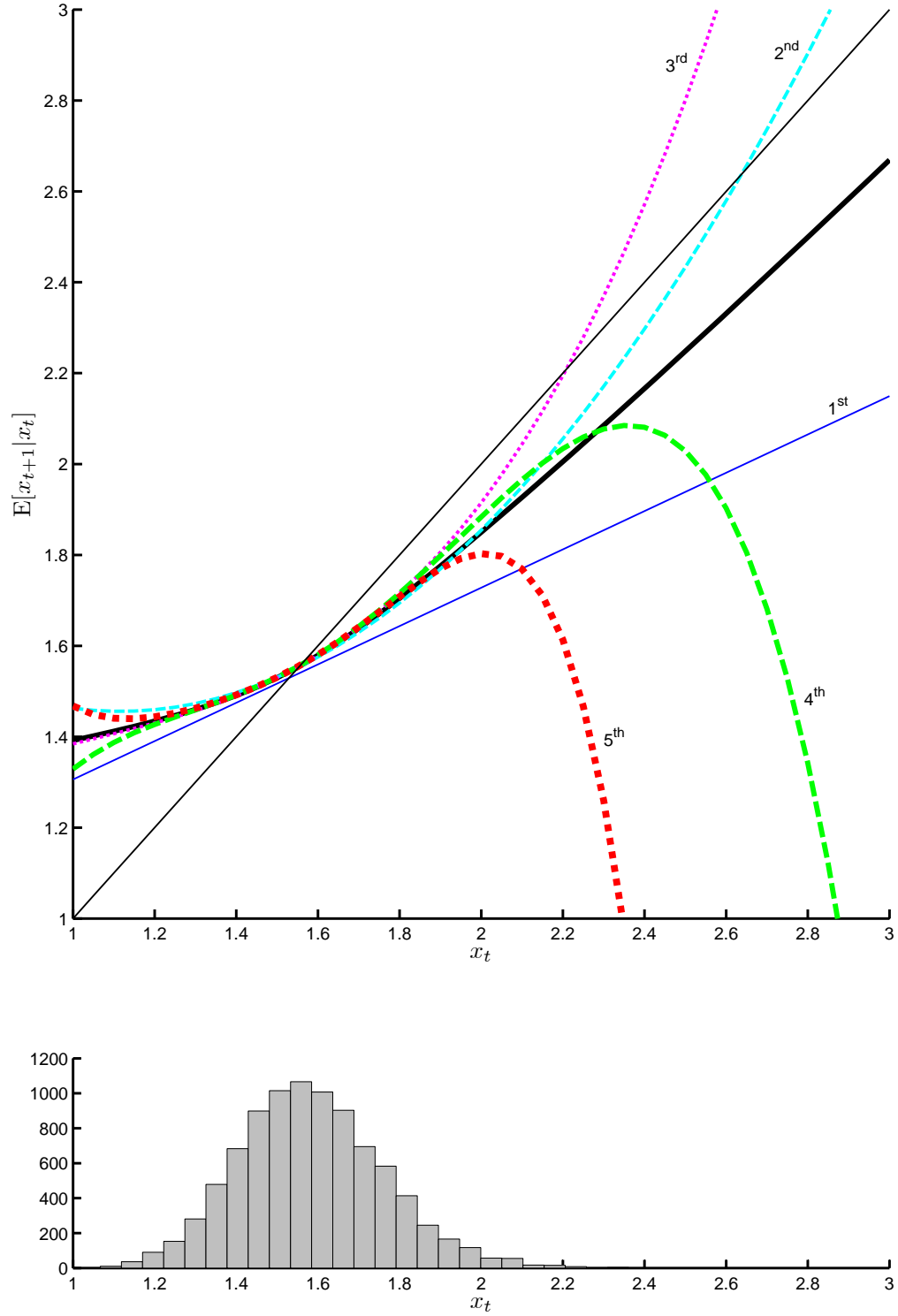


B: Maximum error for 2<sup>nd</sup>-order pruned perturbation



Notes: Panel A (B) plots that part of the generate time series when the two-step ahead perturbation-plus approximation (2<sup>nd</sup>-order pruned perturbation) attains its maximum error.

Figure 10: Higher-order perturbation and the modified Deaton model



Notes: The top panel gives the perturbation approximations and the "truth" (in bold) for the modified Deaton model. The bottom panel gives the distribution for  $x$  according to the "truth".

Table 1: Differences (in %) between approximations and "truth"

Brock-Mirman				
	$\sigma = 0.10$		$\sigma = 0.20$	
	max	mean	max	mean
perturbation + 1	16.7	2.7	25.3	5.4
perturbation + 2	5.2	1.0	8.1	1.9
1 <sup>st</sup> -order pert.	76.1	8.0	142.8	19.2
2 <sup>nd</sup> -order pert.	31.0	1.9	76.4	7.9
2 <sup>nd</sup> -order pruning	47.9	2.0	193.8	8.8
matching				
	$w = 0.96$		$w = 0.97$	
	max	mean	max	mean
"truth"	0.00	0.00	0.00	0.00
perturbation + 1	0.97	0.67	3.09	1.82
perturbation + 1*	0.30	0.12	0.44	0.13
perturbation + 9**	0.66	0.45	2.38	1.38
1 <sup>st</sup> -order pert.	1.02	0.69	3.20	1.87
2 <sup>nd</sup> -order pert.	0.06	0.02	$\infty$	$\infty$
2 <sup>nd</sup> -order pruning	0.36	0.22	1.79	0.94
modified Deaton				
	max	mean		
"truth"	0.1	0.008		
perturbation + 1	97.9	22.1		
perturbation + 2	64.6	11.1		
1 <sup>st</sup> -order pert.	136.9	44.4		
2 <sup>nd</sup> -order pert.	$\infty$	$\infty$		
2 <sup>nd</sup> -order pruning	127.7	12.2		

Notes: "truth" refers to the true policy rule for the Brock-Mirman model and to a very accurate projection method for the other two models. The \* indicates that two separate linear rules are used for the two values of  $z$  and \*\* indicates that the approximation is calculated as described in Appendix C.1. Results are based on a time path of 10,000 observations.

Table 2: Model properties according to approximations and "truth"

	Brock-Mirman							
	$\sigma_z = 0.10$				$\sigma_z = 0.20$			
	E[k]	$\sigma_k$	min	max	E[k]	$\sigma_k$	min	max
truth	0.227	0.123	0.037	1.170	0.334	0.444	0.007	6.868
perturbation + 1	0.218	0.114	0.010	0.983	0.296	0.360	0.000	5.158
perturbation + 2	0.224	0.120	0.029	1.110	0.322	0.417	0.004	6.312
1 <sup>st</sup> -order pert.	0.200	0.099	-0.136	0.552	0.201	0.198	-0.471	0.905
2 <sup>nd</sup> -order pert.	0.223	0.107	0.089	0.838	0.283	0.232	0.089	1.804
2 <sup>nd</sup> -order pruning	0.225	0.108	0.010	0.865	0.299	0.253	0.100	2.155

	matching							
	$w = 0.96$				$w = 0.97$			
	E[n]	$\sigma_n$	min	max	E[n]	$\sigma_n$	min	max
"truth"	0.943	0.0175	0.925	0.961	0.931	0.0315	0.898	0.964
perturbation + 1	0.949	0.0151	0.934	0.965	0.948	0.0218	0.927	0.972
perturbation + 1*	0.941	0.0178	0.924	0.961	0.932	0.0318	0.899	0.966
perturbation + 9**	0.947	0.0157	0.931	0.964	0.944	0.0236	0.920	0.969
1 <sup>st</sup> -order pert.	0.949	0.0151	0.934	0.966	0.949	0.0217	0.928	0.972
2 <sup>nd</sup> -order pert.	0.942	0.0176	0.925	0.961	$\infty$	$\infty$	$\infty$	$\infty$
2 <sup>nd</sup> -order pruning	0.944	0.0155	0.928	0.961	0.937	0.0228	0.915	0.964

	modified Deaton			
	E[a]	$\sigma_a$	min	max
"truth"	0.085	0.098	-0.094	0.848
perturbation + 1	0.057	0.080	-0.101	0.583
perturbation + 2	0.071	0.088	-0.097	0.710
1 <sup>st</sup> -order pert.	0.029	0.070	-0.184	0.344
2 <sup>nd</sup> -order pert.	$\infty$	$\infty$	—	$\infty$
2 <sup>nd</sup> -order pruning	0.079	0.082	-0.043	0.829

Notes: "truth" refers to the true policy rule for the Brock-Mirman model and to a very accurate projection method for the other two models. The \* indicates that two separate linear rules are used for the two values of  $z$  and \*\* indicates that the approximation is calculated as described in Appendix C.1. Results are based on a time path of 10,000 observations.



Table 3: Dynamics for true and incorrect 3<sup>rd</sup>-order pruned perturbation approximation

$t$	$x_t$	$\tilde{x}_t^{(3)}$
1	$\sigma$	$\sigma$
2	$\rho_1\sigma + \rho_2\sigma^2 + \rho_3\sigma^3$	$\rho_1\sigma + \rho_2\sigma^2 + \rho_3\sigma^3$
3	$\begin{aligned} &\rho_1^2\sigma + \rho_1\rho_2\sigma^2 + \rho_1\rho_3\sigma^3 \\ &+ \rho_2\left((\rho_1\sigma)^2 + \mathbf{2\rho_1\rho_2\sigma^3}\right) \\ &+ \rho_3(\rho_1\sigma)^3 \\ &+ O(\sigma^4) \end{aligned}$	$\begin{aligned} &\rho_1^2\sigma + \rho_1\rho_2\sigma^2 + \rho_1\rho_3\sigma^3 \\ &+ \rho_2(\rho_1\sigma)^2 \\ &+ \rho_3(\rho_1\sigma)^3 \end{aligned}$

Notes: The values for  $x_t$  correspond to the values according to the true law of motion, which is given in Equation (71a). The values for  $\tilde{x}_t$  correspond to the values according to the incorrect formulation of the pruning procedure given in Equation (76).