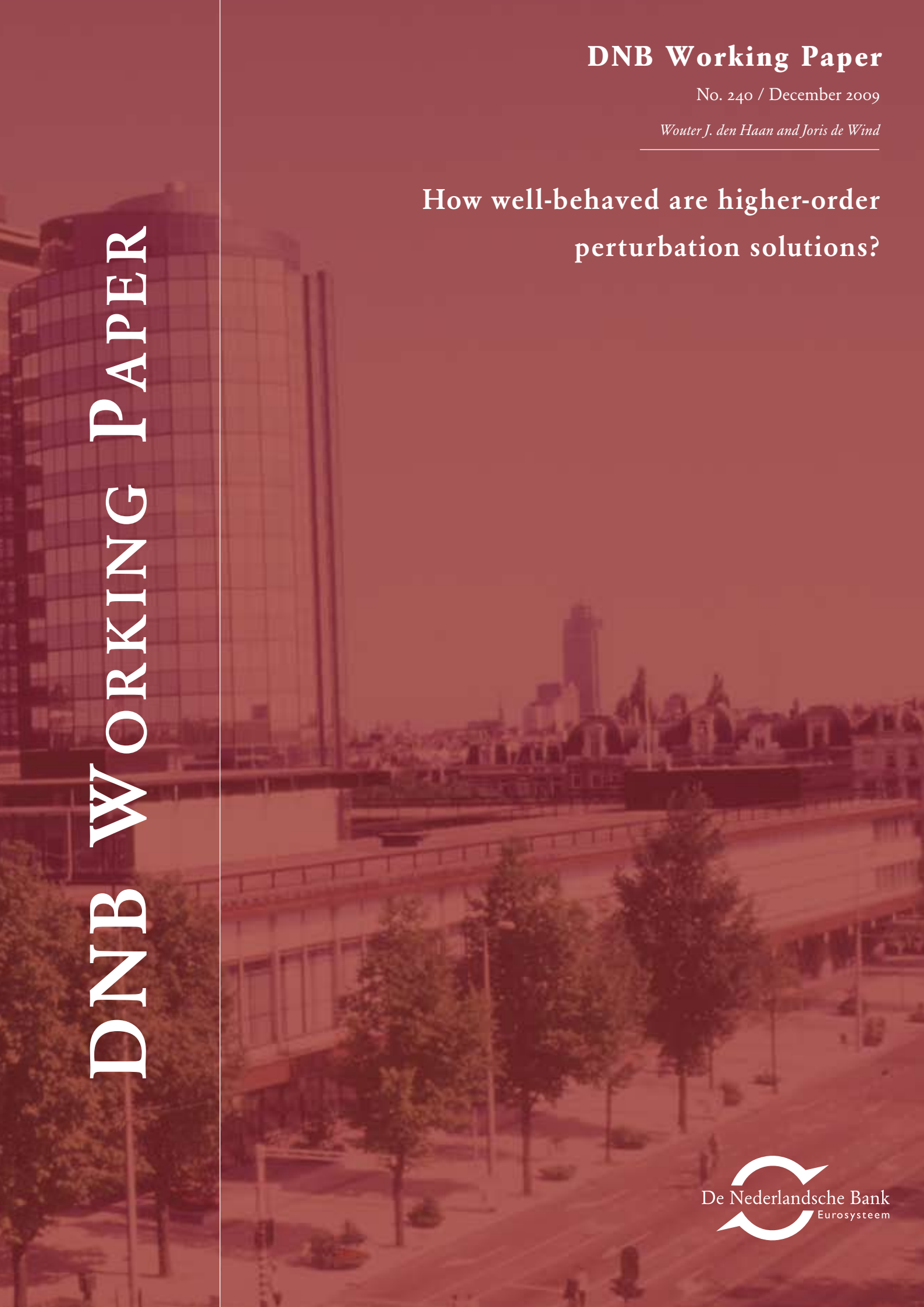
DNB Working Paper

No. 240 / December 2009

Wouter J. den Haan and Joris de Wind

How well-behaved are higher-order perturbation solutions?

DNB Working Paper



De Nederlandsche Bank
Eurosysteem

How well-behaved are higher-order perturbation solutions?

Wouter J. den Haan and Joris de Wind *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 240/2009

December 2009

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

How well-behaved are higher-order perturbation solutions?

Wouter J. den Haan and Joris de Wind *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 240/2009

December 2009

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

How well-behaved are higher-order perturbation solutions?

Wouter J. DEN HAAN and Joris DE WIND*

July 27, 2009

Abstract

They are not well-behaved. The main problem is that one cannot control the radius of convergence when using perturbation techniques. Just outside the radius of convergence, higher-order approximations can easily behave extremely badly, and even within the radius of convergence one can expect higher- but finite-order perturbation solutions to display problematic oscillations. In contrast, with projection methods one can control the radius of convergence. Pruning, the solution proposed to deal with explosive behavior of higher-order perturbation solutions, is shown to be highly distortionary. A simple alternative based on short samples and rejection sampling is proposed and shown to be much less distortive.

Key Words: numerical solutions, perturbations, penalty functions, borrowing constraints

JEL Classification: C63, D52

*den Haan: University of Amsterdam and CEPR, e-mail: wdenhaan@uva.nl. de Wind: De Nederlandsche Bank and University of Amsterdam, e-mail j.de.wind@dnb.nl. We would like to thank Fabio Canova, Michel Juillard, Ondra Kamenik, and Tarik Ocaktan for useful comments.

1 Introduction

Perturbation solution techniques have become a very popular choice to solve dynamic stochastic general equilibrium (DSGE) models. Reasons are the ease with which they can deal with a higher dimensional state space and the development of user friendly software like Dynare and Dynare++ to solve higher-order approximations. This paper highlights some serious shortcomings of higher-order perturbation solutions and questions the appropriateness of these techniques for models in which non-linearities are important. More specifically, this paper makes the following four contributions.

First, we summarize some important results on the radius of convergence of Taylor series expansions. In particular, we are interested in the question in what interval around the perturbation point the Taylor series expansion converges towards the underlying function as the approximation order goes to infinity. We point out that even simple functions like $\ln(x)$ or $\sqrt{x}$ have quite a limited radius of convergence *and* show that just outside the radius of convergence the higher-order Taylor series approximations behave extremely badly. This is obviously not a new result, but given the precipitous increase in the use of this technique it is useful to highlight its limitations with some simple examples.

Second, we show that not only outside, but also within the radius of convergence one is likely to encounter wild oscillations in the numerical solutions for finite-order approximations and convergence towards the truth can be far from monotone. Such oscillations are a feature of all polynomial approximations. The problem with perturbation solutions is particularly troublesome, however, because one cannot control where the oscillations occur; they could be far away from the steady state or close to it. This is an important difference with power series expansions obtained with projection methods with which it is much easier to keep the osciallations outside a certain domain, namely by expanding the grid.

The osciallations of higher-order perturbation solutions can easily lead to explosive time paths. The third contribution of this paper is to document that the pruning procedure, recently proposed by Kim, Kim, Schaumburg, and Sims (2008), creates large systematic distortions and the implied policy rule is not even a function of the model's

state variables.

The fourth contribution of this paper is to consider a very simple alternative that consists of reporting model statistics on the bases of many short samples and using rejection sampling to discard problematic samples. We discuss how initial values can be generated and propose a criterion to discard problematic Monte Carlo samples.

2 Polynomial approximations

In this section, we outline the two main procedures to construct polynomial approximations. It is assumed that one can either calculate the derivatives at one particular point or one can evaluate the function value at a set of nodes.

Polynomial approximations of the function $h(x)$ can be written as

$$p_N(x;\gamma) = \sum_{n=0}^{N} \gamma_n (x-\bar{x})^n, \tag{1}$$

where $\bar{x}$ and the vector γ are the coefficients. One possibility would be to use the Taylor series approximation

$$\begin{aligned} & p_{N,\text{pert}}(x) \\ & = \\ & h(\bar{x}) + \left.\frac{\partial h(x)}{\partial x}\right|_{x=\bar{x}} (x-\bar{x}) + \frac{1}{2!}\left.\frac{\partial^2 h(x)}{\partial x^2}\right|_{x=\bar{x}} (x-\bar{x})^2 + \cdots + \frac{1}{N!}\left.\frac{\partial^N h(x)}{\partial x^N}\right|_{x=\bar{x}} (x-\bar{x})^N. \end{aligned} \tag{2}$$

The main alternative is to use a projection procedure, which requires a grid with J nodes, $x_1, \cdots, x_J$, where $J \geq N$. If one can evaluate the function values at the grid points, then the coefficients of the approximating polynomial can be found by solving the following minimization problem:

$$\gamma = \arg\min_{\tilde{\gamma}} \sum_{j=1}^{J} w(x_j) \left(h(x_j) - p_N(x_j;\tilde{\gamma})\right)^2. \tag{3}$$

If the weights, $w(x_j)$, are equal for all x_j, then this is the same as non-linear least squares.

3 Problems with Taylor series expansions

In Section 3.1, we document that functions commonly used in economics have a limited radius of convergence and that one can expect wild behavior of higher-order approximations outside the radius of convergence. In Section, 3.2, we show how these problems can be dealt with when projection methods are used to solve for the polynomial approximations. In Section 3.3, we show that higher but finite order Taylor series expansion can easily lead to problematic behavior *within* the radius of convergence and make clear that a higher-order approximation could easily be a lot less accurate.

3.1 Simple functions and limited radius of convergence for Taylor series.

Two simple functions that cannot be approximated arbitrarily well with a Taylor series expansion on a relatively small interval around the approximation point are $\ln(x)$ and $\sqrt{x}$. The N^{th}-order Taylor series approximation of $\ln(x)$ around $\bar{x}$ is equal to

$$\begin{aligned} &\ln(\bar{x}) + \frac{\tilde{x}}{\bar{x}} - \frac{1}{2!}\left(\frac{\tilde{x}}{\bar{x}}\right)^2 + \frac{2!}{3!}\left(\frac{\tilde{x}}{\bar{x}}\right)^3 - \frac{3!}{4!}\left(\frac{\tilde{x}}{\bar{x}}\right)^4 + \cdots + (-1)^{N-1}\frac{(N-1)!}{N!}\left(\frac{\tilde{x}}{\bar{x}}\right)^N \\ &\qquad = \\ &\ln(\bar{x}) + \frac{\tilde{x}}{\bar{x}} - \frac{1}{2}\left(\frac{\tilde{x}}{\bar{x}}\right)^2 + \frac{1}{3}\left(\frac{\tilde{x}}{\bar{x}}\right)^3 - \frac{1}{4}\left(\frac{\tilde{x}}{\bar{x}}\right)^4 + \cdots + (-1)^{N-1}\frac{1}{N}\left(\frac{\tilde{x}}{\bar{x}}\right)^N \end{aligned} \tag{4}$$

where $\tilde{x}$ is equal to $x - \bar{x}$. This Taylor series expansion does not converge to $\ln(x)$ if N goes to infinity for $\tilde{x}/\bar{x} > 1$. In fact, the approximation errors diverge to ∞ when $\tilde{x}/\bar{x} > 1$, i.e., when $x > 2\bar{x}$. Interestingly, the Taylor series expansion does converge to $\ln(x)$ for *all* values of x such that $0 < x < \bar{x}$ even though the function value goes to $-\infty$ as x goes to 0.[1] The problems occur when $x > 2\bar{x}$ for which the behavior of $\ln(x)$ seems quite unproblematic. It may in practice, thus, hard to know where problems will occur. The divergence of the Taylor series expansion of $\ln(x)$ is documented in the top panel of Figure 1.A that plots the N^{th}-order Taylor series expansions of $\ln(x)$ for $1 \leq x \leq 2.5$ using $\bar{x} = 1$

[1] The N^{th}-order Taylor series expansion evaluated at x equal to 0, i.e., when $\tilde{x}/\bar{x}$ is equal to -1, is finite but converges to $-\infty$ as N goes to ∞.

and N equal to 1, 2, 5, and 25.[2]

If one uses a first-order approximation, then one obtains a well-behaved approximation, e.g., monotonically increasing like $\ln(x)$ itself, but the approximation is quite bad. If the order of the approximation increases, then the approximation nicely converges to $\ln(x)$ for $x < 2$; using the 5^{th}-order approximation the errors are still noticable for values of x close to 2, but the 25^{th}-order approximation provides a close fit for all values of x in this range.

The cost of using higher-order approximations is that the approximations are highly problematic for $x > 2$. For example, if one uses a second-order, then one looses monotonicity; the approximation is decreasing when $x > 2$. The figure documents that the 5^{th} and 25^{th}-order Taylor series approximation is even more problematic when $x > 2$. When using a 25^{th}-order approximation the approximation basically explodes for values of x just above 2.

In this particular case, there is an easy way to increase the radius of convergence; one can simply increase $\bar{x}$. The bottom panel of Figure 1.A is the analogue of the top panel, but uses $\bar{x} = 1.5$ instead of $\bar{x} = 1$. The figure documents that all Taylor series expansions are well behaved when $x < 2.5$. The drawback of this adjustent is that the approximation errors increase for lower values of x. In particular, the approximation error of the 2^{nd} (5^{th})-order Taylor series expansion at $x = 1$ is equal to 0.0166 (0.0003), which—if $\ln(x)$ is meant to represent percentage deviations—would imply a 1.66% (0.03%) error. Using $\bar{x}$ equal to 1.5 instead of 1, obviously also increases the approximation errors for values of x less than 1, part of the domain not plotted in the figure. In particular, the approximation error of the 2^{nd} (5^{th})-order Taylor series expansion at $x = 0.9$ is equal to 3.08% (0.1%) when $\bar{x} = 1.5$, but is only 0.36% (0.000018%) when $\bar{x} = 1$. Thus, even for the 5^{th}-order approximation serious errors are observed not too far away from the original point of interest, if $\bar{x}$ is equal to 1.5 instead of 1.

Judd (1998) provides a more general discussion. He points out that if a function $h(x)$ has a singularity at x^* that the radius of convergence cannot be larger than $\|\bar{x} - x^*\|$.[3] This

[2] The figures for $\ln(x)$ are more illuminating if the domain is restricted to values above 1; recall that the approximations do converge for values of x less than 1.

[3] See Theorem 6.1.2. A function $h(x)$ has a singularity at x^* if $h(x)$ is analytic on the domain of h

means that one should be careful using Taylor series approximations when the function has singularities. As pointed out by Judd (1998), this may be difficult in economic applications; for example, regularity conditions such as the Inada conditions often imply a singularity. An important degree of freedom one has with any solution procedure and, thus, also when using perturbation techniques, is that one can take transformations of variables, which may expand the radius of convergence and in some cases even eliminate the singularity.[4]

3.2 Difference between Taylor series and projection approximation

The top panel of Figure 1.B plots $\ln(x)$ and the 1^{st}, 2^{nd}, 5^{th}, and 25^{th}-order projection approximation when the Chebyshev nodes used to find the approximation are chosen in the interval $[0, 2]$, the radius of convergence for the Taylor series expansion when $\bar{x} = 1$. The figure documents that when $x > 2$, i.e., when x is outside the grid, that higher-order approximations obtained with projections are also problematic.[5]

The analogue of an increase in $\bar{x}$ for a Taylor series expansion would be an enlargement of the grid. The bottom panel of Figure 1.B plots the results when the nodes are chosen in the interval $[0, 3]$. Again, this pushes the explosive behavior to larger values of x and clearly eliminates the problem for values of x considered here. For lower-order approximations, widening the grid for the projections approximation deteriorates accuracy more than increasing $\bar{x}$ for the Taylor series approximation. In particular, note that the fit of the 1^{st} and 2^{nd}-order approximation is terrible, especially around values of x equal to 1, i.e., around the orginal point of interest.

There is one enormous difference, however, between a Taylor series expansion and

except for x^*.

[4]The analytical solution to the Brock-Mirman version of the standard growth model has a singularity when the capital stock is equal to zero and one can expect this to carry over to other versions of the standard growth model. To deal with this singularity, one could solve for $\tilde{c}(\tilde{k}_{-1}) = \log(c(k_{-1}))$ and $\tilde{k}(\tilde{k}_{-1}) = \log(k(k_{-1}))$ from (for example) $\exp(c) + \exp(k) = \exp(\alpha k_{-1}) + (1-\delta)\exp(k_{-1})$ and $\exp(-c) = \beta \exp(-c_{+1})(\alpha \exp((\alpha-1)k + 1 - \delta)$ instead of solving for $c(k_{-1})$ and $k(k_{-1})$ from $c + k = k_{-1}^{\alpha} + k_{-1}$ and $c^{-1} = \beta c_{+1}^{-1}(\alpha k^{\alpha-1} + 1 - \delta)$.

[5]Note that the 2^{nd}-order projections approximation reaches a peak at 1.8227 and is, thus, not even monotonically increasing within the grid.

a projection approximation. In this particular, example, it turned out to be possible to adjust $\bar{x}$ and increase the radius of convergence for the Taylor series expansion. In practice, however, it either may not be possible or not easy to determine how to adjust $\bar{x}$. With a projection method one can always widen the grid. Moreover, with projection methods, one has more options, for example, one could change the weights of the nodes, increase the number of nodes, or relocate them.

3.3 Local problems with finite-order polynomial approximations

In Section 3.1, we discussed the radius of convergence of Taylor series approximations and the type of behavior one can encounter outside the radius of convergence. The main lesson learned is that when $\bar{x}$ is fixed, the radius of convergence is fixed as well and, thus, cannot be controlled by the researcher. Adjusting $\bar{x}$ may help in terms of changing the radius of convergence, but even if one knows how to adjust $\bar{x}$ to accomplish this, one faces a deterioration of the approximation around the old perturbation point. Projection methods allow for more flexibility and one can, for example, always adjust the location of the grid points.

In this section, we discuss the problems one can encounter with Taylor series approximations in intervals that are within the radius of convergence. In particular, we analyze whether polynomial approximations preserve properties like (i) monotonicity, (ii) having a unique fixed point, and (iii) convergence towards the fixed point, that is,

$$\lim_{t\to\infty} x_t = x^* \ \forall x_0 \text{ if } x_{t+1} = h(x_t) \text{ for } t \geq 0.$$

In the remainder of this section, we focus on the question whether these three properties are easily satisfied for polynomial approximations within the radius of convergence. Section 3.3.1 focuses on the shape of the approximating function and Section 3.3.2 focuses on the convergence towards the fixed point.

3.3.1 Shape of approximation polynomial

Preserving properties like monotonicity is certainly not guaranteed if one uses polynomials.[6] For example, *any* N^{th}-order polynomial $p_N(x)$ will *not* be monotonically increasing for an even value of N. This would not be problematic, if the values of x for which the derivative has the wrong sign are far enough away from the area of interest. The problem with Taylor series expansions is that one cannot control in which part of the domain the approximating function does not preserve monotonicity. Consider the function $h : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ defined as

$$h(x) = 0.5x^{\alpha} + 0.5x, \tag{5}$$

which is monotonically increasing and has a unique positive fixed point at $x = 1$. The top and bottom panel of Figure 2.A plot $h(x)$ together with the second-order Taylor series approximation around x equal to 1 for α equal to 5 and 11, respectively. The second-order approximation is clearly not monotone in the domain of the function and $\partial p_2(x)/\partial x$ is negative for values that are only 30% (11%) below the steady state when α is equal to 5 (11). To obtain monotonicity in a larger interval one could take a Taylor series approximation around a different point. For example, if one takes a Taylor series expansion at x equal to 0.5081 (0.6316) instead of 1 for α equal to 5 (11), then one finds that the second-order Taylor expansion is just monotonically increasing for all values of x in $\mathbb{R}_0^+$, i.e., the domain of the function. But then the error at the original focal point, $\bar{x} = 1$, is equal to 24.2% (41.7%).

Since the function $h(x)$ specified in Equation (5) is a polynomial, one can obtain an exact fit by choosing a high-enough approximation order. But the order of the approximating polynomial has to be quite high. The bottom panel of Figure 2.A also plots the 10^{th}-order Taylor series approximation of $h(x)$ when α is equal to 11, i.e., when $h(x)$ is an 11^{th}-order polynomial. Even when the order of the approximating Taylor series approximation is so close to the order of the approximated polynomial, then the monotonicity is lost in an interval relatively close to the perturbation point.

All second-order polynomial approximations are non-monotonic. Thus, second-order

[6] The same is true for convexity.

approximations obtained with projection methods will, just like the 2^{nd}-order Taylor series approximations plotted in Figure 2.A, be non-monotonic. The difference between Taylor series expansions and power series approximations obtained with projection methods, however, is that one has more control about the shape of the polynomial with projection methods.[7] Figure 2.B plots $h(x)$ for α equal to 5 together with some second-order approximations obtained with projections. The second-order approximations are obtained by OLS using a relatively low, i.e., three, number and a relatively high, i.e., eleven, number of Chebyshev nodes on the interval $[0, 1.5]$. While the second-order approximation obtained with only 3 nodes is not monotonic over the domain of the function, the second-order approximation obtained with 11 nodes is. Thus, with projection methods it is fairly easy to find a 2^{nd}-order approximation that is and monotonically increasing over this interval and much more accurate then the 2^{nd}-order Taylor series expansion around $x = 1$.

3.3.2 Explosive behavior

If the approximating function does not preserve key properties of the true function, then it is obviously also not clear whether it will preserve the poperty that iterating on the function will generate a sequence that converges to the fixed point. An advantage of Taylor series approximations is that the true fixed point is always a fixed point of the approximating functions as long as the Taylor series approximation is around the true fixed point. This does not mean, however, that you will always converge towards it. In fact, the Taylor series approximation may have additional fixed points, even if the approximated function has a unique fixed point.

Consider the function

$$h(x) = \alpha_0 + x + \alpha_1 e^{-\alpha_2 x}. \tag{6}$$

The value of α_0 is chosen such that $x = 1$ is a fixed point and the value of α_1 is chosen to ensure that the function is always increasing for $x \geq 0$. Figure 3 plots the function and the second-order approximation when $\alpha_0 = -0.3495$, $\alpha_1 = 0.95$, and $\alpha_2 = 1$. The graph

[7] For example, if one would locate all grid points close to the boundaries, then the higher-order terms get little weight and one basically guarantees monotonicity.

shows that the second-order approximation has a fixed point at $x = 1$, but also a second fixed point, x^{**}, at $x = 3$. If the Blanchard-Kahn conditions are satisfied and there is, thus, a locally unique fixed point, then *all* second-order perturbation approximations will have an additional fixed point.

Moreover, if $x_0 > 3$, then the function values would diverge. But x^{**} is quite a bit higher than the true fixed point and as long as the initial value of x is not that high, then the generated sequence still converges towards the true fixed point. It is easy to show, however, that the additional fixed point of the second-order Taylor series approximation is equal to

$$x^{**} = 1 + \frac{2}{\alpha_2},$$

which means that as α_2 increases the additional fixed point gets closer to the original fixed point.

But no matter the value of α_2, one can always ensure convergence by making sure that the initial value of x is not too high. If we add a stochastic shock to the analysis, it may not be always possible to avoid explosive behavior. For example, suppose that the true policy function is given by

$$h(x, \theta) = -0.9 + \theta + x + 0.9e^{-x}, \tag{7}$$

where θ is a stochastic variable. Its mean is set equal to $0.9(1-e^{-1})$ so that the steady state value of x is still equal to 1. Its support is assumed to be the interval $[0, \zeta]$ with $\zeta < 0.9$, which ensures that the support of x is also a compact set. Since the function is linear in the stochastic variable, uncertainty does not affect the perturbation approximations.

The second-order approximation of $h(x, \theta)$ is equal to

$$p_2(x, \theta) = 1 + \theta - 0.9(1 - e^{-1}) + 0.6689(x - 1) + 0.149(x - 1)^2,$$

which—for low enough values of θ—has a stable fixed point equal to the true fixed point and for a value of x above the stable fixed point also an unstable fixed point. When

$$\theta > \bar{\theta} = 0.9 - 0.45e^{-1},$$

then the function has no fixed points. The value of $\bar{\theta}$ is 29% above the steady state value. In models with idiosyncratic shocks, this would not be an unreasonable high realization.[8] Suppose that for some time θ is above $\bar{\theta}$ and there is, thus, no fixed point and, thus, no upper bound on x. The value of x would then start to increase. It is possible that by the time θ has reduced to a value at which there is again a fixed point for x, the value of x has become so high that it is to the right of the second unstable fixed point and the values of x would continue to explode even though θ has returned to values at which the second-order approximation has a unique locally stable fixed point.

In the next section, we use a DSGE model to document, however, that this quite easily does happen, that is, the simulated series often explode.

4 Approximate solutions to DSGE models

Finding approximations to the policy functions that solve DSGE models is more difficult than the approximation problem discussed in Section 2, because policy functions are only implicitly defined. That is, one cannot directly evaluate the function $h(\cdot)$ at a set of grid points nor calculate its derivatives directly.

A standard DSGE problem can be written as follows:

$$\mathrm{E}\left[f(k_{-1}, \theta, \theta_{+1}, h(k_{-1}, \theta; \Lambda), \Lambda | k_{-1}, \theta\right] = 0 \quad \forall k_{-1}, \theta, \tag{8}$$

where $f(\cdot)$ is a known function, k_{-1} a vector of endogenous state variables determined in the last period and known at the beginning of the current period, Λ is a vector of structural parameters, and $h(k, \theta; \Lambda)$ is the unknown policy function, that depends on the state variables and the structural parameters Λ. The exogenous random state variable, θ, has a known law of motion.[9]

[8] Suppose the log of the random variable follows an AR(1), the innovation variation is equal to 0.015, and the autoregressive coefficient is equal to 0.95. These are values that are sensible for *aggregate* random variables. The standard deviation is 4.8% and even if we would consider a four standard deviation shock, then we are still far away from the 29%. But for idiosyncratic random variables, the variance of the random variables can easily be a multiple of those appropriate for aggregate random variables.

[9] To simplify the notation, we assume that $h(\cdot)$ and θ are scalars, but the analysis does not depend on

For the standard growth model, we would have the following two equations:

$$\mathrm{E}\left[\begin{array}{c} -\frac{1}{\theta k_{-1}^{\alpha}+(1-\delta)k_{-1}-h(k_{-1},\theta)} \\ +\beta\frac{\alpha\theta_{+1}h(k_{-1},\theta)^{\alpha-1}+1-\delta}{\theta_{+1}h(k_{-1},\theta)^{\alpha}+(1-\delta)h(k_{-1},\theta)-h(h(k_{-1},\theta),\theta_{+1})} \end{array}\middle| k_{-1},\theta\right]=0, \text{ and} \tag{9}$$

$$\ln(\theta_{+1})=\rho\ln(\theta)+\varepsilon_{+1} \quad \varepsilon_{+1}\sim N(0,\sigma_{\varepsilon}^{2}). \tag{10}$$

In the remainder of this section we discuss the two types of procedures with which one can obtain polynomial approximations of the unknown policy functions.

4.1 Projection methods

For a given approximation, $p(k_{-1},\theta;\gamma)$, and a procedure to numerically calculate the conditional expectation, one can calculate for each element of the state space the Euler equation error defined as

$$u(k_{-1},\theta)=\mathrm{E}\left[\begin{array}{c} -\frac{1}{\theta k_{-1}^{\alpha}+(1-\delta)k_{-1}-p(k_{-1},\theta;\gamma)} \\ +\beta\frac{\alpha\theta_{+1}p(k_{-1},\theta;\gamma)^{\alpha-1}+1-\delta}{\theta_{+1}p(k_{-1},\theta;\gamma)^{\alpha}+(1-\delta)p(k_{-1},\theta;\gamma)-p(p(k_{-1},\theta;\gamma),\theta_{+1};\gamma)} \end{array}\middle| k,\theta\right] \tag{11}$$

Projection methods solve for the coefficients γ by constructing a grid for k_{-1} and θ and then minimizing some loss function over the Euler equation errors on the grid.

The theory on the convergence of approximations obtained with projection methods is well established if one can evaluate a function $h(x)$ at a set of grid points. In this case, polynomial approximations converge *uniformly* under weak conditions *if* one uses Chebyshev nodes.[10] When solving for the policy function of a DSGE model, however, the function $h(x)$ is only implicitly defined and one cannot directly evaluate $h(x)$ for given values of the state variables. Even for simple iterative procedures, however, convergence to the true policy function has been established in the literature for several models.[11]

this assumption.

[10] See Judd (1992) and Section 6.7 in Judd (1998).

[11] Examples are Coleman (1991), Marcet and Marshall (1994), and Rendahl (2006). Rendahl (2006) use splines to interpolate, but the convergence result also holds for polynomial approximations, since polynomial interpolation leads to uniform convergence as long as Chebyshev nodes are used.

4.2 Perturbation methods

The perturbation procedure introduces a scalar variable, σ, that controls the amount of uncertainty.[12] This method finds the coefficients of the approximating polynomials, $p_N(k_{-1}, \theta, \sigma; \gamma)$, by sequential differentiation of Equation (9) and evaluation of the expresssions at $\sigma = 0$ and the steady state values of k_{-1} and θ. There are two parts to the question whether $p_N(k_{-1}, \theta, \sigma; \gamma)$ converges to the true solution $h(k_{-1}, \theta; \Lambda)$.

1. The first question is whether this procedure backs out the correct derivatives of $h(k_{-1}, \theta; \Lambda)$ evaluated at the steady state. The implicit function theorem can be used to show that the true derivatives of $h(k_{-1}, \theta; \Lambda)$ can indeed be obtained by sequential differentiation of Equation (9).[13]

2. The second question is whether the function $h(k_{-1}, \theta; \Lambda)$ can be approximated arbitrarily well with its Taylor series approximation in a large enough interval to make the approximation meaningful. This is a question that has received very little attention. Section 3, however, documented that the radius of convergence of a Taylor series approximation can be small even for a simple well-behaved functions and that even within the radius of convergence higher- but finite-order approximations can be problematic.

5 A DSGE example

In Section 3, we used simple examples to highlight the problems one may encounter when using Taylor series approximations. In this section, we show that one can easily encounter these problems in a simple DSGE model. We use the model of Deaton (1991) in which agents face idiosyncratic income shocks and use one-period bonds to smooth consumption.

[12] In the problem specified in Equation (9), there is only one source of uncertainty and σ is equal to σ_ε. To simplify the notation, we suppress the other structural parameters as an argument of $p_N(\cdot)$.

[13] See Section 13.1 in Judd (1998). Since the stochastic variable in Equation (9) has continuous support, it does not fit exactly into the conditions of the standard implicit function theorem used in Judd (1998); see Jin and Judd (2002) for a discussion on perturbation solutions when random processes do have continuous support.

To be able to use polynomial approximations, we replace the borrowing constraint used in Deaton (1991) with a penalty function.[14]

5.1 Model

The agent's maximization problem is given by

$$\max_{\{c_t,a_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \beta^{t-1} \frac{c_t^{1-\nu}-1}{1-\nu} - P(a_t) \tag{12}$$

$$\text{s.t.}$$

$$c_t + \frac{a_t}{1+r} = a_{t-1} + \theta_t,$$

$$\theta_t = \bar{\theta} + \varepsilon_t \text{ and } \varepsilon_t \sim N(0,\sigma^2),$$

$$a_0 \text{ given.}$$

Here c_t is the agent's consumption level, a_t the amount of assets chosen in period t, and θ_t the random and exogenous income level. The penalty function, $P(a_t)$, is a function that is decreasing in a_t and, thus, makes choosing low asset holdings costly.

Penalty function. Deaton (1991) uses a non-negativity constraint on a to restrict borrowing, which would correspond to the following penalty function:

$$P(a) = \begin{cases} \infty & \text{if } a < 0 \\ 0 & \text{if } a \geq 0 \end{cases} \tag{13}$$

We use the following penalty function:

$$P(a) = \frac{\eta_1}{\eta_0} \exp(-\eta_0 a) - \eta_2 a. \tag{14}$$

One reason behind choosing this particular functional form is that it does not have a single singularity and it, thus, avoids the problems with a limited radius of convergence discussed

[14] See den Haan and de Wind (2009) for a further discussion on using penalty functions instead of inequality constraints in this type of model.

in Section 3.1.[15,16] In fact, if any non-polynomial function should be unproblematic for perturbation solutions, then it should be the exponential.

Parameter values. The parameter values used are chosen to ensure that the inequality constraint has a non-trivial impact in the original model.[17] In particular, each period the inequality constraint is binding for 28% of all agents. When solving the model with the penalty function, we use the same parameter values, but replace the inequality constraint with the penalty function given in Equation (14). We consider two values for the curvature parameter, η_0, namely 10 and 20. As η_0 increases the penalty function gets closer to the penalty function implied by the inequality constraint, but the non-linearity will make the model more difficult to solve. We choose the other two parameter values, η_1 and η_2, such that both the mean and the standard deviation of a_t in the model with the penalty function are identical to the values found in the model with the inequality constraint.[18]

In den Haan and de Wind (2009), we discuss in detail how the properties of the model with a smooth penalty functions differ from the properties of the model with an inequality constraint, both when an extremely accurate solution is used to solve the model and when perturbation approximations are used. Here we interpret the two models (with a low and a high value of η_0) simply as two non-linear models without non-differentiabilities that are close to an important model considered in the literature and discuss the difficulties one may run into when solving this model with perturbation solutions.

[15] Obviously, using a penalty with an unlimited radius of convergence does not imply that the same holds for the policy function, but if one starts with a penalty function that has itself a limited radius of convergence, then this is unlikely to be helpful in reducing the radius of convergence for the policy function.

[16] Kim, Kollmann, and Kim (2009) and Preston and Roca (2006) also solve a model with a penalty function with perturbation methods, but their penalty functions do have a singularity.

[17] They are as follows: $\beta = 0.9$, $r = 0.03$, $\nu = 3$, $\bar{\theta} = 1.5$, and $\sigma = 0.15$.

[18] In finding these parameter values, we use an extremely accurate solution procedure to solve both models and we are confident that the outcome is not affected by approximation errors. Section A documents the accuracy of our projections solution.

5.2 First-order conditions and policy functions

The first-order conditions of the agent's maximization problem are given by

$$\frac{c_t^{-\nu}}{1+r} + \frac{\partial P(a_t)}{\partial a_t} = \beta \mathrm{E}_t \left[c_{t+1}^{-\nu} \right] \tag{15}$$

and the budget constraint.[19] Since we assume that the agent's income, θ_t, is i.i.d., and the interest rate is constant, there is only one state variable, namely cash-on-hand, x_t, which is equal to $a_{t-1} + \theta_t$. The objective, thus, is to find policy functions, $a(x_t)$ and $c(x_t)$ that satisfy the first-order condition and the budget constraint.

Perturbation solution. Since asset holdings can be negative, we obtain a perturbation solution in the level of a. The solution for consumption is obtained directly from the budget constraint. Given that there is no analytical solution to this model, there is no guarantee that there is no singularity in the policy functions $c(x_t)$ and $a(x_t)$, but we think that there is none. For example, the policy functions are unproblematic when x_t is equal to zero or negative. The agent would consume a positive amount and increase his debt level. This in contrast with the standard growth model in which consumption and savings would be zero when the capital stock is equal to zero and these functions would not be well defined for negative values of the capital stock.

5.3 Approximations of the policy function

Figures 4.A and 4.B plot the perturbation solutions (1^{st} through 5^{th} order) together with a very accurate projections solution for η_0 equal to 10 and 20, respectively.[20] The figure documents that the solutions are close to each other around the steady state, but that they are very different at points in the state space that are further away. To evaluate the relevance of the points in the state space where the differences occur, we plot in the bottom panel of each figure the histogram of the state variable according to the accurate

[19] By letting the penalty be a utility penalty, the budget constraint is identical to the budget constraint of the model with an inequality constraint.

[20] In Section A we document the accuracy of our "accurate" projections solution and the lack of accuracy for the perturbation solutions.

projections solution as observed in a sample of 9,000 observations. One should keep in mind, however, that the distribution generated by the perturbation solutions could very well be different than the one plotted in the figure, which is generated with the accurate projections solution. For example, small deviations of a numerical solution in the tail can push the simulated series into an area where the inaccuracies of the numerical solution are larger. In fact, for some perturbation solutions the simulated data explode and the implied support would, thus, be unbounded.

The maximum value of the state variable, x, observed in our long simulation generated with a very accurate solution is equal to 2.35 (2.36) when η_0 is equal to 10 (20) and the corresponding minimum value is equal to 1.01 (1.02). By plotting the functions for values of x between 1 and 2.5, we go slightly beyond these extremums.

Figure 4.A documents that the higher-order perturbation solutions are indistinguishable from the accurate solution in a large part of the relevant state space. Difference widen away from the steady state and interestingly, the 2^{nd} and 4^{th}-order approximation remain closer to the accurate solution than the 3^{rd} and 5^{th}-order approximation.

The higher-order approximations solutions considered have much stronger non-linearities than the accurate solution. In fact, the highest-order perturbation approximation considered, i.e., the 5^{th}-order seems most problematic. It starts deviating from the accurate solution when x takes on a value that is around the 98^{th} percentile and the policy function that has been gradually increasing up to this point, then starts decreasing rapidly.

Finally, note that the 2^{nd} and 3^{rd}-order approximation are a bit more convex than the accurate solution, but nicely preserve its shape. In contrast, the 4^{th} and the 5^{th}-order approximation are not monotonically increasing and do not preserve the property of convexity satisfied by the accurate solution. Nevertheless, the 2^{nd} and 3^{th}-order perturbation solutions turn out to be a lot more problematic when the approximations are used to generate simulated data.

5.4 Location of oscillations and uncertainty

As the amount of uncertainty increases, the simulated series are more likely to enter the area with the problematic oscillations when the numerical solution is *not* adjusted for the higher level of uncertainty. Higher-order perturbation solutions are—in contrast to the first-order solution—affected by the amount of uncertainty. So the question arises, whether an increase in uncertainty pushes the problematic oscillations of the approximating functions further away from the steady state. Figure 5 plot the 5^{th}-order polynomial when σ_ε is equal to 0 and 0.3. According to the perturbation solution, the main effect of an increase in σ_ε on the solution is an upward shift. Although the osciallations still occur at the same place the problems are actually more severe for two reasons. First, since the standard deviation has increased, the solution is more likely to reach the problematic area. One would want the osciallations to occur further away from the steady state to compensate for this, but this does not happen. Second, the upward shift implies that the average state variable moves away from the steady state making it also more likely that the problematic region is reached.

5.5 Underlying non-linearity

Figures 6.A and 6.B plot the Taylor series approximation of the penalty *term*, that is, the derivative of the penalty function, which shows up in the first-order condition. To make the graphs comparable with Figures 4.A and 4.B that plot the policy function, we plot the penalty term as a function of the state variable, x, not as a function of the chosen level of a. That is, for a given value of x, we use the accurate projections solution to calculate the asset choice, a, and then evaluate the Taylor series expansion of $\partial P(a)/\partial a$ at this choice. The lower panel plots again the histogram of the state variable.

The graph makes clear that the penalty term is not approximated well with Taylor series approximations of 5^{th} or lower order, even though, it does not have a singularity and can be approximated globally with a Taylor series expansion.

In our simple model it is very clear which part of the model contains the non-linearities and if one is limited to solve the model with perturbations, then one could consider alter-

native penalty functions that are easier to approximate.[21] In practice, the non-linearities of a DSGE policy function are likely to be generated by a combination of factors. Nevertheless, it may be a worthwhile exercise to ask oneself whether the components of the model, such as the utility function, adjustment costs specification, production function, are approximated well with a low-order Taylor series expansion if one uses perturbation analysis to obtain a numerical solution.

5.6 Approximations and model properties

Table 1 reports key properties of the model when data are generated with the different numerical solutions. Besides the five perturbation solutions we also consider a very accurate projection solution.[22]

When η_0 is equal to 10, then the model properties as generated by higher-order perturbation solutions are quite close to those generated by the accurate projections solution. The largest differences are obtained for the behavior in the tail. For example, the maximum value of a observed in our sample of 9,000 observations is 0.66, 0.59, and 0.54 for the 3^{rd}, 4^{th}, and 5^{th}-order perturbation solution and for the accurate projections solution we found this value to be equal to 0.59.

Explosive 2^{nd} and 3^{rd}-order solution. When η_0 is equal to 20, then the data simulated with the 2^{nd} and 3^{rd}-order perturbation solution exploded; consequently no summary statistics are reported for these numerical solutions. According to Figure 4.B, the 2^{nd} and 3^{rd}-order perturbation solutions are closer to the accurate solution than the 4^{th} and 5^{th}-order perturbation solution in several aspects, in particular, they preserve the monotonicity and convexity of the approximated function. The problem with the 2^{nd} and 3^{rd}-order perturbation solution is that they are *steeper* than the true solution and as the

[21] In particular, den Haan and de Wind (2009) discuss the advantages of directly specifying a polynomial for the penalty term of the same order as the perturbation order used.

[22] All numerical solutions are for the model with the smooth penalty function. See den Haan and de Wind (2009) for a discussion on whether models with a not-too-nonlinear penalty function are very different from the model with a non-negativity constraint for a.

numerical approximation becomes steeper, it runs the risk of getting into the explosive region.

Non-explosive but strange looking 4^{th} and 5^{th}-order solution. Although, the data generated with the 4^{th} and 5^{th}-order perturbation solution did not explode, these numerical approximations are far from satisfactory. Several important statistics generated with these two numerical solutions are quite different from the accurate solution. For example, the standard deviation of a is equal to 0.097 according to the accurate solution, but equal to 0.111 and 0.087 according to the 4^{th} and 5^{th}-order perturbation solution, respectively. Moreover, the statistics generated by the 5^{th}-order perturbation solution are not always closer to those generated with the accurate solution than the statistics generated by the 4^{th}-order perturbation solution. For example, the correlation between the change in consumption and the change in assets is equal to 0.895 according to the accurate solution, but equal to 0.852 and 0.811 according to the 4^{th} and 5^{th}-order perturbation solution, respectively. This is consistent with Figure 4.B that documents that the 5^{th}-order perturbation solution is less accurate than the 4^{th}-order solution for many relevant values of x.

CDFs according to the different solutions. Figure 7 plots (part of) the CDFs according to the different numerical solutions. The outcome is again that there are substantial differences between the different numerical solutions.[23] The difference between the CDF of the accurate approximation and the CDF according to the 4^{th}-order perturbation solution is mainly quantitatively, but they have a similar shape. The shape of the CDF according to the 5^{th}-order solution, however, is very different. It has a very unusual shape in that it does not converge gradually towards 100%, but approaches 100% quickly. This means that in the distribution of a there is substantial mass close to the maximum value. This explains why the observed maximum value in our sample of simulated data according to the 5^{th}-order approximation is much lower than the maximum value according to

[23] We only report the results for $\eta_0 = 20$, because for $\eta_0 = 10$, the CDFs are quite close to each other.

the 4^{th}-order perturbation or the accurate projections procedure.[24] This odd distribution generated by the 5^{th}-order perturbation solution is due to the policy function being humpshaped as documented in Figure 4.B.

6 Avoiding explosive behavior

In this section, we discuss two procedures to deal with the explosive behavior that one can easily encounter if one uses higher-order perturbation solutions. The first is the pruning procedure proposed by Kim, Kim, Schaumburg, and Sims (2008). We find this procedure quite problematic and instead propose a procedure that is based on using multiple short samples.

6.1 Pruning

Section 6.1.1 describes the pruning procedure, Sections 6.1.2 and 6.1.3 document that the pruning procedure distorts the underlying numerical solution quite a bit, and 6.1.4 compares the model properties when the pruning procedure is and when it is not used to modify the numerical solution.

For the results in this section, it does not make much difference whether a numerical solution is accurate or not. The focus of this section is not whether a particular solution is correct but what effects pruning has.

6.1.1 Pruning procedure

Kim, Kim, Schaumburg, and Sims (2008) describe a procedure that guarantees that the data simulated with a higher-order perturbation solution is stationary. In this section, we describe the procedure and its problems. The N^{th}-order perturbation solution can be

[24] The maximum according to the 5^{th}-order solution is 0.30, whereas the maximum according to the 4^{th}-order perturbation solution and the accurate projections solution are equal to 0.58 and 0.64, respectively.

written as

$$p_{N,\text{pert}}(a_{t-1},\theta_t) - \bar{a}_N \tag{16}$$

$$=$$

$$\gamma_{N,k}\left(a_{t-1}-\bar{a}_N\right)+\gamma_{N,\theta}\left(\theta_t-\bar{\theta}\right)+\tilde{p}_{N,\text{pert}}(a_{t-1}-\bar{a}_N,\theta_t-\bar{\theta}),$$

where a_{t-1} is the endogenous state variable, θ_t is the exogenous random variable, $\bar{a}_N$ is the *stochastic* steady state and $\tilde{p}_{N,\text{pert}}(\bar{a}_{t-1}-\bar{a}_N,\theta_t-\bar{\theta})$ is the non-linear part of the perturbation solution.[25] For our DSGE example, with i.i.d. income shocks, one only needs one state variable, $a_{t-1}+\theta_t$, but it may be useful to describe the procedure for the more general case.

The pruning procedure consists of the following steps.

1. Simulate a_t^* using

$$a_t^* - \bar{a}_N = \gamma_{N,k}\left(a_{t-1}^* - \bar{a}_N\right) + \gamma_{N,\theta}\left(\theta_t - \bar{\theta}\right). \tag{17}$$

   Kim, Kim, Schaumburg, and Sims (2008) actually propose to simulate a_t^* using

$$a_t^* - \bar{a}_1 = \gamma_{1,k}\left(a_{t-1}^* - \bar{a}_1\right) + \gamma_{1,\theta}\left(\theta_t - \bar{\theta}\right), \tag{18}$$

   that is, the first-order perturbation solution. The advantage of using Equation (17) to generate a_t^* is that it is closer to the numerical solution of interest specified in Equation (16). For example, the steady state value of a according to Equation (16) is equal to the steady state value according to Equation (17), but not to the value according to Equation (18). The important feature of a_t^* is that it is stationary whenever the Blanchard-Kahn conditions are satisfied.[26]

[25] The stochastic steady state is the value of a that satisfies $p_{N,\text{pert}}(a,\bar{\theta})=a$. For $N=1$, the stochastic steady state is equal to the non-stochastic steady state. The stochastic steady state of θ_t is equal to the non-stochastic steady state, namely $\bar{\theta}$, because its law of motion is linear. There is no constant term in Equation (16), since the variables are expressed relative to the stochastic steady state.

[26] It is possible, but unlikely, that the slope coefficients of (18) guarantee stationarity, but the coefficients of (17), that are affected by the amount of uncertainty in the problem, do not. In this case one could simply use $a_t^* - \bar{a}_N = \gamma_{1,k}\left(a_{t-1}-\bar{a}_N\right)+\gamma_{1,\theta}\left(\theta_t-\bar{\theta}\right)$, which in our opinion is still better than using (18).

2. Simulate $a_{\text{prune},t}$ using

$$a_{\text{prune},t} - \bar{a}_N \tag{19}$$
$$= \gamma_{N,k}\left(a_{\text{prune},t-1} - \bar{a}_N\right) + \gamma_{N,\theta}\left(\theta_t - \bar{\theta}\right) + \tilde{p}_{N,\text{pert}}(a^*_{t-1} - \bar{a}_N, \theta_t - \bar{\theta}).$$

Since a^*_t and θ_t are stationary so is $\tilde{p}_{N,\text{pert}}(a^*_{t-1} - \bar{a}_N, \theta_t - \bar{\theta})$. Consequently, $a_{\text{prune},t}$ is stationary.

In the remainder of this section, we document that this procedure is quite distortive.

6.1.2 Pruning and the speed of convergence function

In describing the following example, we let a_t refer to the data simulated by the 2^{nd}-order perturbation solution, $a_{\text{prune},t}$ refer to the 2^{nd}-order perturbation solution when pruning is used, and a^*_t to the (modified) first-order solution given in Equation (17).

Understanding what pruning does is often not very easy, since one has to understand the dynamics of an auxilliary state variable that is moving separately from the actual state variable but still affects it. Key in understanding the distortion introduced by pruning is the shape of the underlying function, here the 2^{nd}-order perturbation solution, and the (modified) first-order approximation. Figure 8 plots these two numerical solutions.[27] For values of the state variable below the steady state, the second-order approximation lies above the first-order approximation, which means that convergence towards the steady state goes faster if the second-order approximation is used. The opposite is true when a_{-1} is above its steady state value.

Figure 9 plots the convergence paths when $a_1 = a_{\text{prune},1} = a^*_1$, that is, when at the beginning of period 2, the inputs are the same for all three policy rules. In the top panel, the initial value of a is equal to 0.5 and in the bottom panel it is equal to -0.5. In addition to the convergence paths according to the first and the second-order approximation the figure also plots the convergence path when pruning is used. To understand the results,

[27] The policy function in our model is a function of cash-on-hand, $x = a_{-1} + \theta$. In drawing this graph, however, income is set equal to its unconditional mean and we plot the policy function as a function of beginning-of-period asset holdings, a_{-1}.

it is important to remember that the pruning solution has two inputs, namely $a_{\text{prune},t-1}$ and a^*_{t-1}.

First consider the case when the initial value of a_{-1} is equal to -0.5, which is illustrated in the bottom panel. Although the pruning policy rule is a mix of the first and the second-order approximation, convergence is actually fastest when pruning is used. The reason is the following. For $t = 2$, the pruning solution and the choice according to the second-order solution are identical, that is, $a_2 = a_{\text{prune},2}$, because both of the inputs used in the pruned 2^{nd}-order policy rule are identical to the input of the not-pruned 2^{nd}-order policy rule. The choice according to the first-order solution is—as was point out above—below the choice of the second-order solution, that is, $a^*_2 < a_2$. For $t = 3$, the choice generated by the not-distorted second-order perturbation solution is based only on a_2. In contrast, for the pruning solution the linear part is also based on $a_{\text{prune},2} = a_2$, but the non-linear part is based on a^*_2. Given that $a^*_2 < a_2$, the non-linear part of the pruning solution is based on a value for a that is further away from the steady state. Given the convexitiy in the policy function this implies a larger adjustment.

Next, consider the case when the initial value of a_{-1} is equal to 0.5. For initial values above the steady state, the 1^{st}-order solution converges faster to the steady state, which means that the non-linear part of the pruned 2^{nd}-order solution converges quickly as well.

This exercise illustrates that pruning distorts the solution, but the qualitative results are not that different. The following exercise reveals better the extend to which the pruning solution distorts the underlying policy function.

6.1.3 The implied policy "function" of the pruning procedure

The pruned numerical solution generates time paths for $\{a_{\text{prune},t}\}_{t=1}^{T}$ and $\{x_{\text{prune},t}\}_{t=1}^{T}$, where $x_{\text{prune},t} = a_{\text{prune},t-1} + \theta_t$. Given that there is only one state variable in our model, we can evaluate what according to the simulated data the pruning policy rule for a_t looks like as a function of the state variable x_t.[28]

[28] In contrast, the underlying policy rule is—as is clear from Equation 18— a function of two state variables, namely $x_{\text{prune},t} = a_{\text{prune},t-1} + \theta_t$ and $x^*_{t-1} = a^*_{t-1} + \theta_t$.

Figures 10.A and 10.B plot these implied policy "functions". The figures also plots the results when the actual, i.e., not distorted, perturbation solution is used. The striking observation is that the approximations based on the pruning procedure are no longer functions. That is, when the pruning procedure is used, different choices are generated for identical state variables.

Figure 10.A plots the results for the 2^{nd} and 3^{rd}-order perturbation solution. We can make the following observations. First, the pruned policy rule is *not* a function. Second, for low values of x the pruned policy rule predicts values that are systematically above the underlying policy function and for high values of x the bias has the opposite sign. The bias is especially sever for high values of x. This is not surprising given that for the high values of x the underlying policy function is too steep and the solution runs the risk of explosion. So although, there should be some bias in this part of the state space, the picture makes clear that pruning also distorts the policy function for low values of x.

Figure 10.B plots the results for the 4^{th} and 5^{th}-order perturbation solution. We find again that for low values of the state variable the pruned choices turn out to be systematically above the not-distorted second-order solution and for high values they are systematically below the not-distorted second-order solution. The bottom panel for the 5^{th}-order perturbation solution makes clear that pruning does not eliminate the erroneous oscillation in the underlying 5^{th}-order perturbation solution. But pruning is also not intended to do so, since the fact that the policy function starts to decrease for larger values of the state variable does not induce explosive behavior; in fact, it only makes the function more stable.

The following example, documented in Figure 11, makes clear why the pruned approximation is not a function. It also shows that, although the distortion is stronger if one is further away from the steady state, the pruned policy rule is not a function close to the state either, unless, of course, the underlying approximation is linear.

Figure 11 plots a generated time path according to the 1^{st}, 2^{nd}, and pruned 2^{nd}-order perturbation solution. All procedures start using the same level of savings in the beginning of period 2, thus, $a_1 = a_{\text{prune},1} = a_1^*$ and we set this initial value equal to 0.4. This is point

"A" in the graph. The initial value is above the steady state, so if θ_2 would be equal to its steady state value, then the asset holdings would decrease towards the steady state value. We set θ_2 equal to its steady state value and we choose θ_3 so that the pruned solution in period three, i.e., point "B" in the graph, is equal to the initial value, that is,

$$a_{\text{prune},3} = a_{\text{prune},1}. \tag{20}$$

Given that we are above the steady state, this means that θ_3 has to be above $\bar{\theta}$ to make a increase back up again. In period 4, we set θ_4 equal to $\bar{\theta}$, as it was in period 1. Thus, by construction we have

$$x_{\text{prune},4} = a_{\text{prune},3} + \bar{\theta} = x_{\text{prune},2} = a_{\text{prune},1} + \bar{\theta}. \tag{21}$$

Thus, the only state variable in the model, x, is the same in periods 2 and 4 according to the data generated with the pruned solution. But note that

$$a_{\text{prune},4} < a_{\text{prune},2}, \tag{22}$$

that is, different choices are made in these two periods, even though the state variable is identical. The intuition is the following. In period 2, the pruned choice of the second-order policy rule and the choice that is not pruned are identical. The reason is that in period 2, the linear as well as the non-linear part of the pruned decision is based on the same value, that is, $x_{\text{prune},2} = x_2^*$. In period 4, the linear part of the pruned choice is based on $x_{\text{prune},4} = a_{\text{prune},3} + \bar{\theta}$, corresponding to point "B" in the graph, which has the same value $x_{\text{prune},2}$. The non-linear part of the pruned choice, however, is based on $x_4^* = a_3^* + \bar{\theta}$. This corresponds to point "C" in the graph and lies below point "B". Consequently, the pruned choice in period 4 lies below the pruned choice in period 2, even though the beginning-of-period values of cash-on-hand are identical in both periods.

6.1.4 Pruning and model properties

The previous subsection documented that pruning distorts the approximating policy functions. Consequently, model properties based on the pruning procedure will be different too. If the numerical solution without pruning explodes, then one cannot measure the

impact of pruning on the numerical solution, but for the other cases one can. Table 2 is the equivalent of Table 1 for the case when the pruning procedure is applied. We also apply the procedure to the accurate solution.[29]

Figures 4.A and 4.B document that for a wide range of values the different numerical solutions are quite similar to each other, but that when they differ they differ a lot. If pruning would avoid the simulated series to take on values in the tail, then one could expect that after pruning the numerical solutions are closer to each other. This is true in the sense that with pruning none of the series explode. But comparing the generated model properties then we find after pruning a similar disparity as before pruning. For example, the pruned 5^{th}-order perturbation solution has a ridiculously low maximum value, just like the not-pruned perturbation solution, due to the erroneous hump in this perturbation solution.

The top panel of Figure 12 plots (part of) the CDFs for the data generated with the different pruned solutions. A comparison between figure 7 and 12 makes again clear that pruning does affect the numerical solution quite a bit. For example, the CDF of the 4^{th}-order numerical solution lies substantially below the CDF of the pruned 4^{th}-order solution, even though the 4^{th}-order solution is not even explosive.

Considering the pruned accurate solution is most instructive to understand the distortive effects of pruning. Our accurate solution is a very well-behaved policy rule and we never encountered explosive behavior no matter how long the sample. Moreover, as documented in the appendix this numerical solution is extremely accurate and one can expect the true model properties to be very close to the ones generated with this numerical solution, at least, if pruning is not used. In terms of the implied policy function, the pruned accurate policy function generates a systematic bias similar to the ones documented in Figure 10.A for the second-order perturbation solution.

A comparison of Tables 1 and 2 documents that pruning results in very different model properties. For example, whereas the mean level of asset holdings is equal to 0.0849 when

[29] The non-linear part is defined as the accurate solution less the standard first-order perturbation solution.

pruning is not used ($\eta_0 = 20$), this drops to 0.0752 when pruning is used. Similarly, the standard deviation drops from 0.0973 to 0.0844. Given that the not-pruned numerical solution is shown to be accurate, the pruned solution is not and given the magnitude of the differences it is very inaccurate. Even when η_0 is equal to 10 and the non-linearities are substantially smaller, we find that the model properties are affected a lot by pruning.

6.2 Averaging short samples

Section 6.2.2 describes an alternative procedure that avoids the problems of pruning and Section 6.2.3 discusses the results when this alternative procedure is used for the Deaton model. But we start with a short discussion on how to report model properties.

6.2.1 Background

There are two common ways used in the literature to report model statistics. The first is to report unconditional moments, typically approximated with the sample moments of a very long sample of simulated data. The second is to report the average of the statistic across Monte Carlo replications, where in each replication a sample with typical length is used, say 160 observations for a quarterly model. The classic Kydland and Prescott (1982) paper uses the second procedure. The second procedure not only provides a value for the statistic of interest, but also accounts for small sample bias and provides a measure of the variation across Monte Carlo replications. For example, it may be possible that the statistic measured with actual data is not that close to the average across Monte Carlo replications, but is still well within the range of generated values.

For first-order moments the results of the two procedure are identical except for some remaining sampling noise. For higher-order moments the results are typically not identical, even if one uses an extremely long sample in the first procedure and zillions of Monte Carlo replications. For example, if

$$x_t = 0.9x_{t-1} + \varepsilon_t, \;\; \varepsilon_t \sim N(0,1), \tag{23}$$

then the unconditional variance of x_t is equal to

$$\mathrm{E}\left[\left(x_t - \mathrm{E}\left[x_t\right]\right)^2\right] = \mathrm{E}\left[x_t^2\right] = 5.26, \tag{24}$$

whereas the mean across Monte Carlo replications of the sample variance

$$\frac{\sum_{t=1}^{T}\left(x_t - \left(\sum_{t=1}^{T} x_t\right)/T\right)^2}{T}$$

is equal to

$$\mathrm{E}\left[\frac{\sum_{t=1}^{T}\left(x_t - \left(\sum_{t=1}^{T} x_t\right)/T\right)^2}{T}\right] = \mathrm{E}\left[x_t^2\right] - \mathrm{E}\left[\left(\frac{\sum_{t=1}^{T} x_t}{T}\right)^2\right] < \mathrm{E}\left[x_t^2\right]. \tag{25}$$

The second procedure produces a smaller variance, because in each of the short samples the mean of x_t is adjusted.

Researchers may have a preference for one of the two procedures, but both are valid and sensible ways to produce model statistics. We argue that with a little bit of care the second procedure can be used to diminish the problems of explosive behavior.

6.2.2 Short-sample procedure

There are two reasons why calculating statistics with the small sample procedure is helpful in avoiding distorting the underlying function. Then the Monte Carlo replication does not have to be rejected and no distortion takes place. First, because the samples are shorter, the simulation may not reach the problematic area. Second, eliminating some Monte Carlo replications should not distort the results very much as long as the fraction of discarded samples is not too high. If one does end up discarding many samples, then one either has a very bad numerical approximation (most likely case) or one really does have a very non-linear underlying policy function. In both cases one should very seriously consider using an alternative to perturabtion procedures.

Care should be given in how to design the short-sample procedure. In Monte Carlo studies, the initial conditions of the short sample cannot be the steady state, because this could bias the results in non-linear models. Typically a long sample is used to obtain

randomized initial conditions. But if one first has to run a long sample, then the simulated series may still explode taking away the advantage of the short-sample procedure. Also, one would need a criterion to decide whether a particular Monte Carlo replication has to be discarded. In the remainder of this section, we explain how we propose to deal with both issues.

Procedure to draw initial observation. Our Monte Carlo procedure uses initial values that are distributed according to the ergodic distribution of the observations used in the small samples, i.e., the one that is based on a rejection criterion. This procedure starts with an intial value, for example, the stochastic steady state. It then simulates T observations where T is the length of the Monte Carlo sample. If this sample is not rejected according to the rejection criterion discussed below then one uses the last observation as the initial observation for the next Monte Carlo replication. If the sample is rejected, then one draws a new sample. This procedure needs a number of Monte Carlo replications to reach the ergodic distribution. We used 500 Monte Carlo replications to initialize the procedure and then use the next 1,000 Monte Carlo replications to actually calculate statistics. The results reported in the tables are based on this procedure.[30]

Selection criterion. Before simulating any short sample, first generate one very long sample of the state variables and calculate the mean, μ_a, and the variance, σ_a^2, using the 1^{st}-order perturbation solution. Recall that with the first-order solution the simulated series are by construction stationary.

The rejection procedure consists of the following step. Given the data of a Monte Carlo replication calculate

$$\hat{\sigma}_{a,T}^2 = \frac{\sum_{t=1}^{T} \left(x_t - \mu_a\right)^2}{T}. \tag{26}$$

Note that $\hat{\sigma}_{a,T}^2$ is calculated using μ_a the mean of a observed in a very long simulation generated with the first-order solution and is *not* calculated with the sample mean of a_t in the short sample. This makes it much more straightforward to compare $\hat{\sigma}_{a,T}^2$ with σ_a^2.

[30] An alternative would be to draw the initial conditions from a long simulation generated with a first-order solution.

In fact, we have[31]

$$\mathrm{E}\left[\hat{\sigma}_{a,T}^2\right] = \sigma_a^2. \tag{27}$$

The idea is to discard a short-sample if $\hat{\sigma}_{a,T}^2$ is "too large" relative to σ_a^2. In particular, a Monte Carlo replication is discarded if

$$\hat{\sigma}_{a,T}^2 > \Lambda^2 \sigma_a^2.$$

The short-sample procedure makes sense if (i) the value of Λ is not too high and (ii) the fraction of discarded samples is low, say less than 5%. Note that the standard deviation of a according to the accurate solution is only 39% above the value according to the first-order solution when $\eta_0 = 20$. One should keep in mind that this is found for a very long sample. The sample length of our Monte Carlo replications is equal to 160 so a doubling of the standard deviation may not be impossible even if the order of the approximation is not increased. Nevertheless we start with $\Lambda = 2$ and consider 3 as an alternative value. We would be hesitant to go to a much higher value, because if the numerical solution is that non-linear one seriously should ask oneself whether perturbation is the right solution procedure.[32]

The advantage of this procedure, is that the not rejected samples are not distortive at all, whereas the pruning procedure always modifies the simulated data even when it does not explode. Of course, our procedure could be distortive if one ends up rejecting many samples, but in our opinion one should keep Λ fairly low *and* only use the results when the rejection rate is low. Note that there are two reasons why a MC sample is rejected: (i) the results are affected by the type of erroneous oscillations highlighted in this paper that are not part of the true policy rule and one should discard or (ii) the results are affected by sharp non-linearities that are part of the true solution and one should not throw out. If the rejection rate is high because of the second reason then one should increase Λ, but we still want to argue that one should be careful in doing so.

[31] The reason we use the variance and not the standard deviation is that this equality would not hold for the standard deviation.

[32] It is also insightful to know why the sample variance is so high. Is it because the series are truly exploding or just occasionally taking on values far away from the mean.

Suppose that (i) in the short-sample procedure one finds a high rejection rate and (ii) one has a numerical solution that leads to explosive solutions when used in a long simulation and one believes this should not be part of the true solution. Note that it makes sense to increase Λ if the numerical solution is such that it is actually quite accurate, but just happens to lead to explosive behavior when it is not accurate. But given that it is much more likely that the explosive behavior is due to erroneous oscillations in the numerical solution, we think that one should not consider values of Λ that are too high.

6.2.3 Short-sample procedure and model properties

Table 3 reports the results when Λ is equal to 2. This means that a Monte Carlo replication is rejected if the variance is 4 (2^2) times as high as the variance of the first-order solution. When η_0 is equal to 10 then none of the 1,000 Monte Carlo replications is rejected. This is, of course, the ideal case. One can simply report the results being sure that there is no distortion to the policy rule at all. When η_0 is equal to 20, however, then a substantial number of replications is rejected. There are three possibilities: (i) because the sample is short the variance can easily be higher than in the long sample even though we calculate both using the same mean, (ii) the true policy function generates much more volatility than the first-order solution, or (iii) the numerical solutions are so bad that one even runs into wild behavior in short samples. Since we have an accurate solution as the stand in for the truth we know that the true policy function does indeed generate occasionally a variance in the short sample that is more than four times as high as the variance of the linear solution. Note that 7.8% of the solutions of the accurate solution are rejected. Below we discuss the distortion introduced by these rejections that for the accurate solution should not have occured.

As discussed above a value of Λ equal to 2 may be conservative for the sample length considered here. As documented in Table 4, when Λ is equal to 3 then all rejection rates are low except those for the 3^{rd}-order perturbation solution. Instead of raising Λ or thinking of another way to deal with the wild behavior, it seems better not completely ignore a perturbation solution that requires such a high value of Λ. The chance is much higher

that the explosive behavior is due to erroneous oscillations then that one has an almost accurate solution that just occasionally leads to explosive behavior.

Does this procedure distort the outcomes. First consider the outcomes for our accurate projection solution, referred to as "truth" in the tables. When we compare the statistics based on one long sample with those based on short samples and $\Lambda = 3$, then there are some differences. But these are *not* due to distortions, because when Λ is equal to 3, then not a single Monte Carlo replication is rejected. These differences are due to the fact that the average of higher-order small sample statistics is not equal to the unconditional value of this statistic. The bottom panel of Figure 12 plots the CDFs of the short-sample data generated with the different numerical solution for the case when η_0 is equal to 20 and the value of Λ is set equal to 3. A comparison with Figure 7 makes clear that there is little distortion, except of course, that now the results for the 2^{nd} and 3^{rd}-order approximation do not explode and can be reported.

To understand whether this procedure is distortive one should compare the results when Λ is equal to 2 and some Monte Carlo samples are rejected with the results when Λ is equal to 3 when there are no rejections. In this case, we know that one would be overly conservative if one would use Λ equal to 2 instead of 3. But how much is the mistake one makes by discarding 7.8% of the samples. The mean drops from 0.0851 to 0.0835 a drop of almost 2%. Clearly a non-trivial error but minuscule relative to the distortions observed when the pruning solution is used. For the other statistics the differences are even smaller so even with Λ is equal to 2 one would have obtained a quite accurate set of statistics.

Now consider the 4^{th}-order perturbation solution for which the number of rejections drops from 43.3% to 0.1% if Λ increases from 2 to 3. Obviously, if the number of rejections drops by so much, then there is a good chance that the value of Λ used was too low. But note that even in this case the impact on the reported statistics is relatively minor. The mean value of a is 8.3% lower when Λ is equal to 2, but the other statistics are quite similar.

7 Concluding comments

This paper has highlighted the problems one can expect when using perturbation solutions, both when the state variables are outside and when they are inside the radius of convergence. Although these problems are not systematically discussed in the literature, many researchers have faced the consequences in terms of explosive behavior of simulated series.

It is extremely unlikely that numerical solutions that generate explosive solutions would pass a serious accuracy test and one wonders why researchers had enough confidence in their numerical solution to use it to simulate data. But it is a sad sign of modern economics that accuracy tests are out of fashion and numerous results based on numerical approximations have been published without a proper discussion on the accuracy of the numerical solution.

Of course, it is possible that a solution is accurate in most of the state space and that the consequences of the inaccuracies are such that—even though they are unlikely to happen often—have far reaching consequences, for example, because the numerical solution would then start generating a diverging path of values.

In those cases, it is worth considering a modification of the numerical solution that avoids these problems, but we want to stress that it does not make sense to use such a modification unless one has convinced oneself that the problematic explosive behavior is due to unlikely events and the solution is accurate for most values of the state variables.

The paper has shown that the modification proposed in the literature is highly distortive and introduces a systematic bias even when the policy function does not have to be modified to guarantee stability. As an alternative, we propose to use a small-sample procedure. This procedure is much less distortive, because it only is implemented when the simulated series do behave wildly.

A Accuracy

Table 5 reports the Euler equation errors calculated on a grid with 15,001 equidistant nodes on $[1, 2.5]$ and using 30 Gaussian quadrature nodes to calculate the conditional expectation. The errors of our accurate projections procedure are very small, which is the main purpose of this exercise. The errors for the perturbation solutions are quite bad, but this do not add any new information above the direct comparison of the policy functions, given the accuracy of the projections solution.

References

Coleman, W. J. I. (1991): "Equilibrium in a Production Economy with an Income Tax," *Econometrica*, 59, 1091–1104.

Deaton, A. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59, 1221–1248.

den Haan, W. J., and J. de Wind (2009): "Incorporating Limits on Borrowing in DSGE Models: Problems and Solutions for Numerical Solution Procedures," Unpublished manuscript, University of Amsterdam.

Jin, H.-H., and K. L. Judd (2002): "Perturbation Methods for General Dynamic Stochastic Models," unpublished manuscript, Stanford University.

Judd, K. L. (1992): "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory*, 58, 410–452.

——— (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.

Kim, S., R. Kollmann, and J. Kim (2009): "Solving the Incomplete Markets Model with Aggregate Uncertainty Using a Perturbation Method," *Journal of Economic Dynamics and Control*, forthcoming.

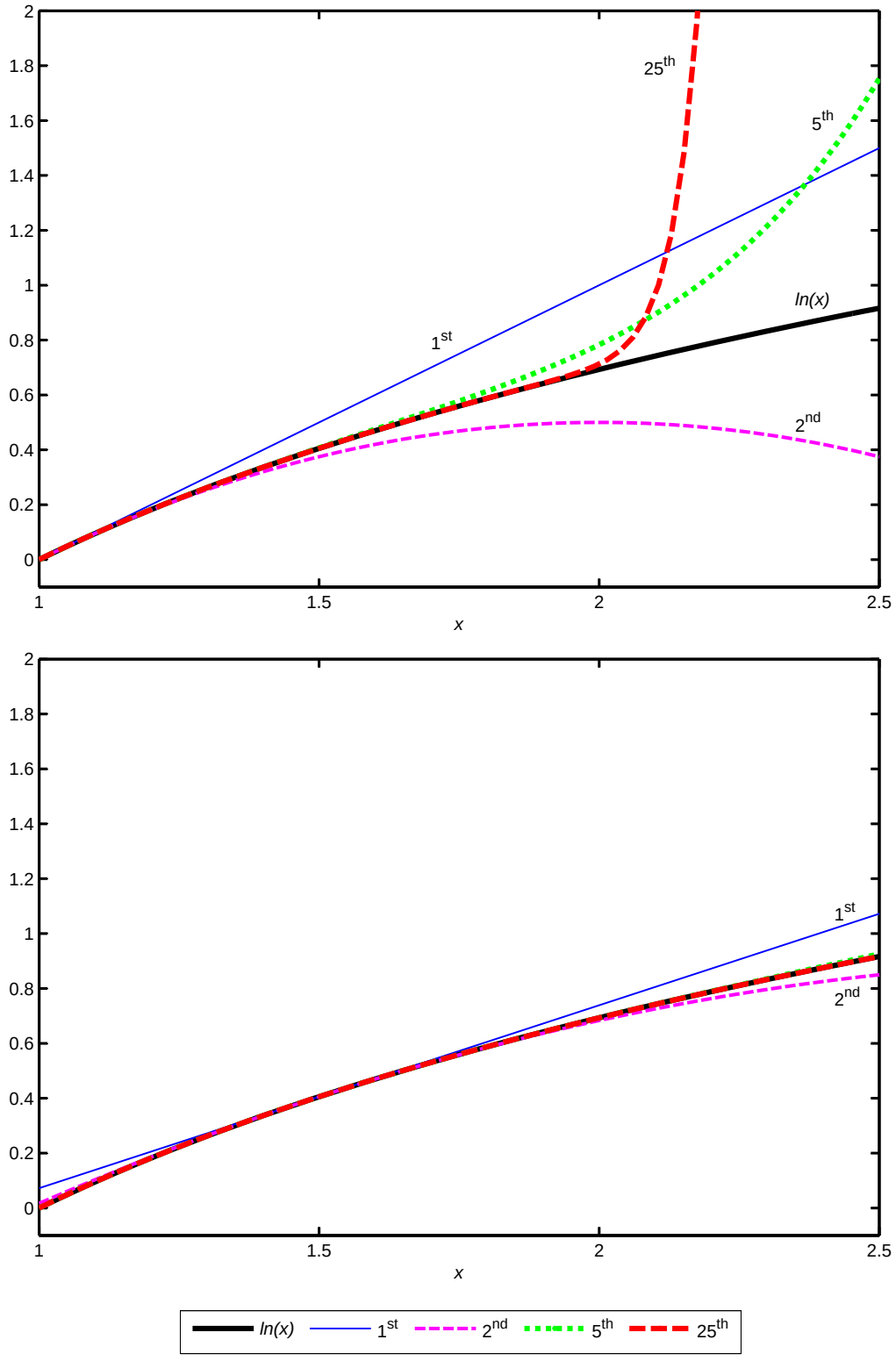
Kydland, F., and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345–1370.

Marcet, A., and D. Marshall (1994): "Solving Nonlinear Rational Expectations Models by Parameterizing Expectations: Convergence Results in the Stationary Case," Unpublished manuscript, Universitat Pompeu Fabra.

Preston, B., and M. Roca (2006): "Incomplete Markets, Heterogeneity and Macroeconomic Dynamics," unpublished manuscript, Columbia University.

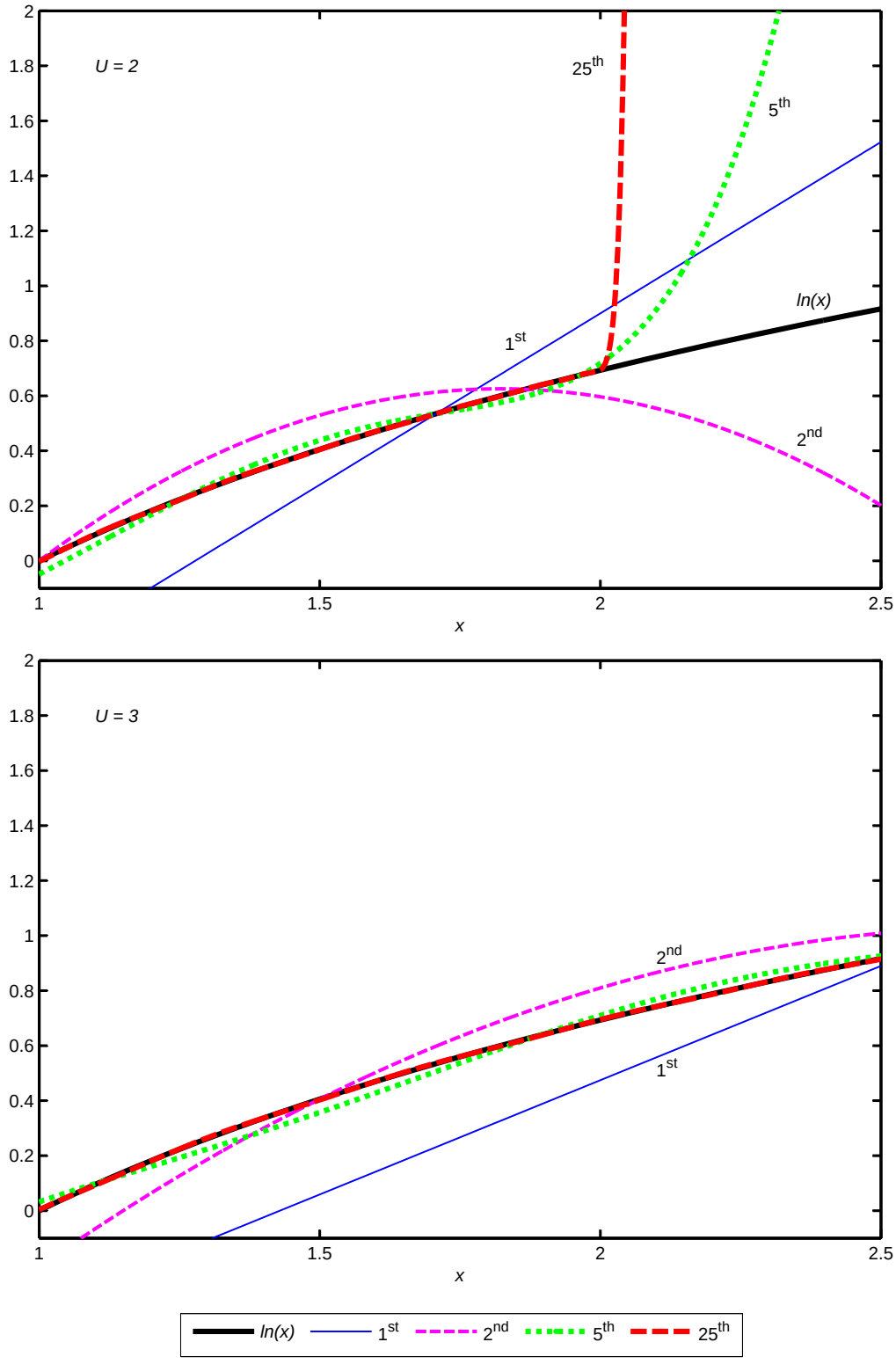
Rendahl, P. (2006): "Inequality Constraints in Recursive Economies," Unpublished manuscript, University of California at Davis.

Figure 1.A: $ln(x)$ and its Taylor series approximation around $\bar{x}$



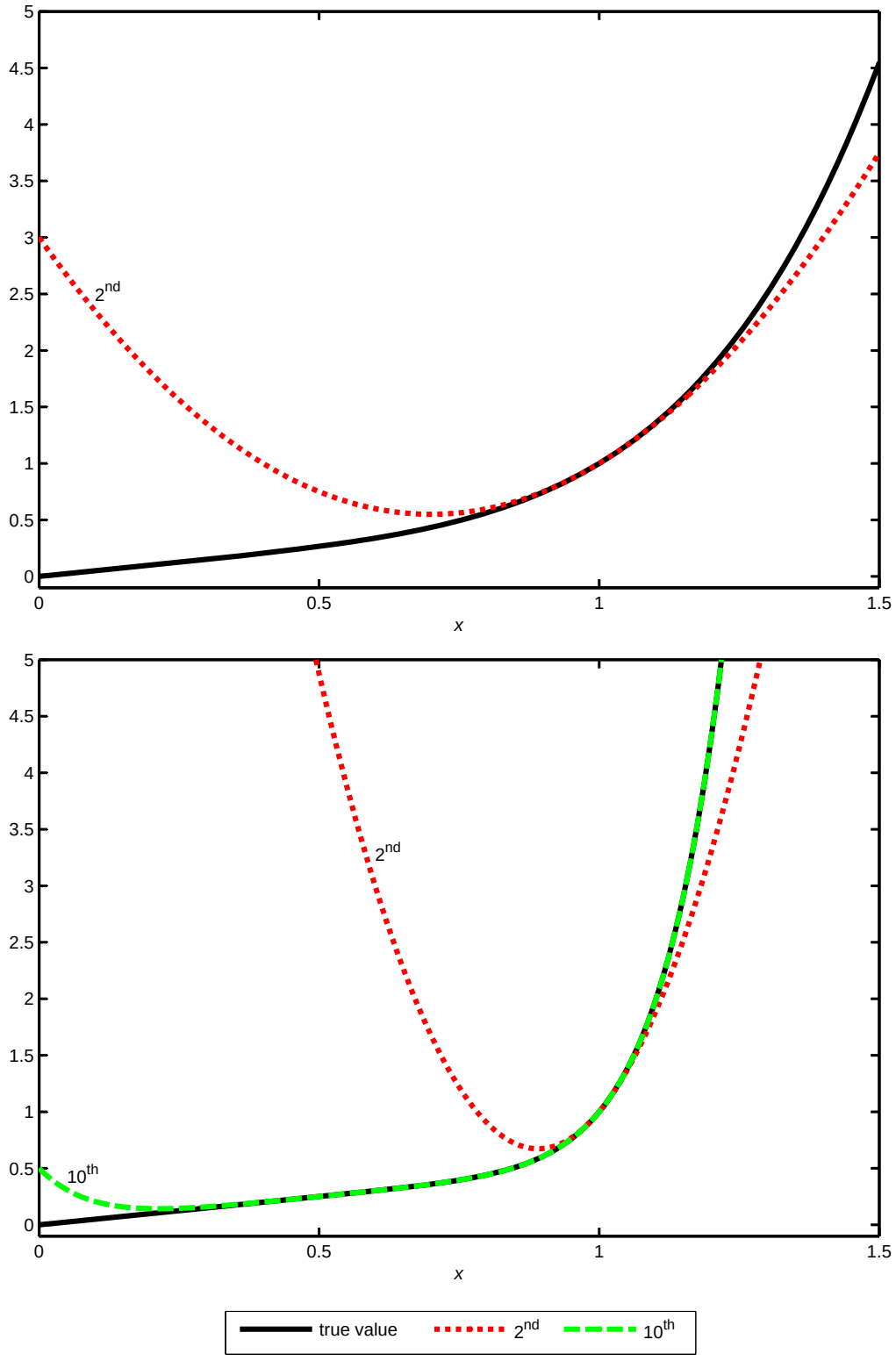
Notes: The graphs plot $ln(x)$ and its Taylor series expansion at the indicated order. The value of $\bar{x}$ is equal to 1 in the top panel and equal to 2 in the bottom panel.

Figure 1.B: $ln(x)$ and projections polynomial approximations; nodes in $[0, U]$



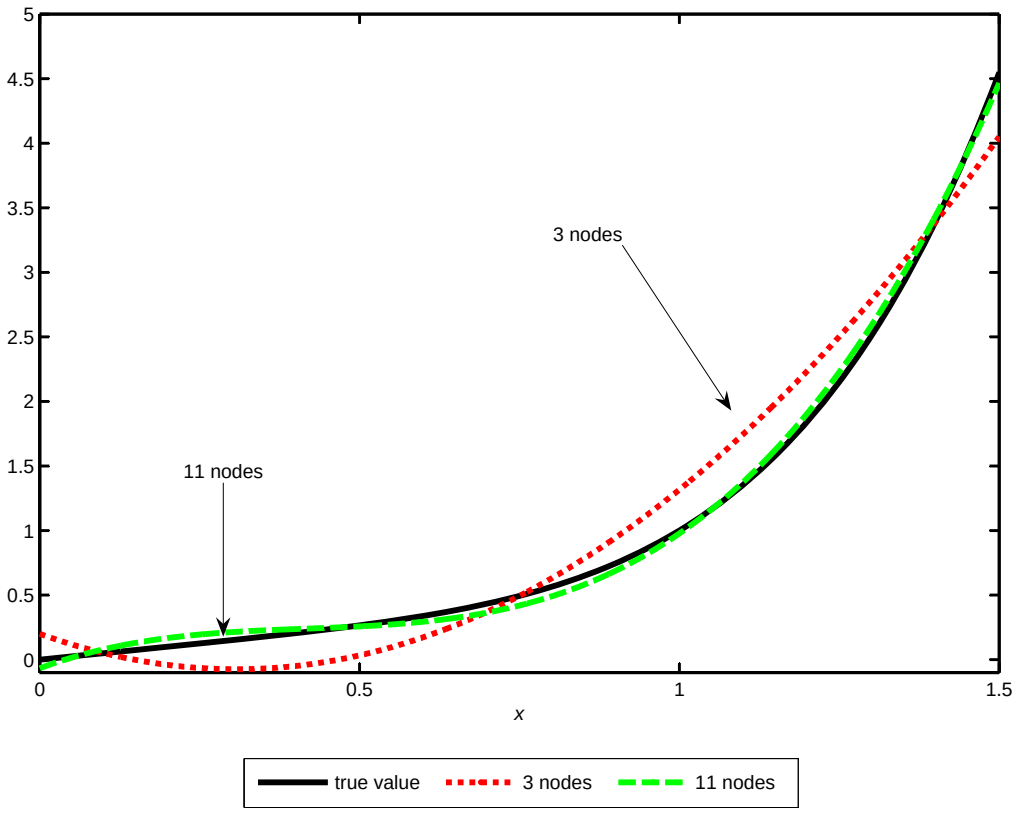
Notes: Chebyshev points are used to locate the nodes

Figure 2.A: $x^{\alpha} + x$ and its Taylor series approximations around $x = 1$



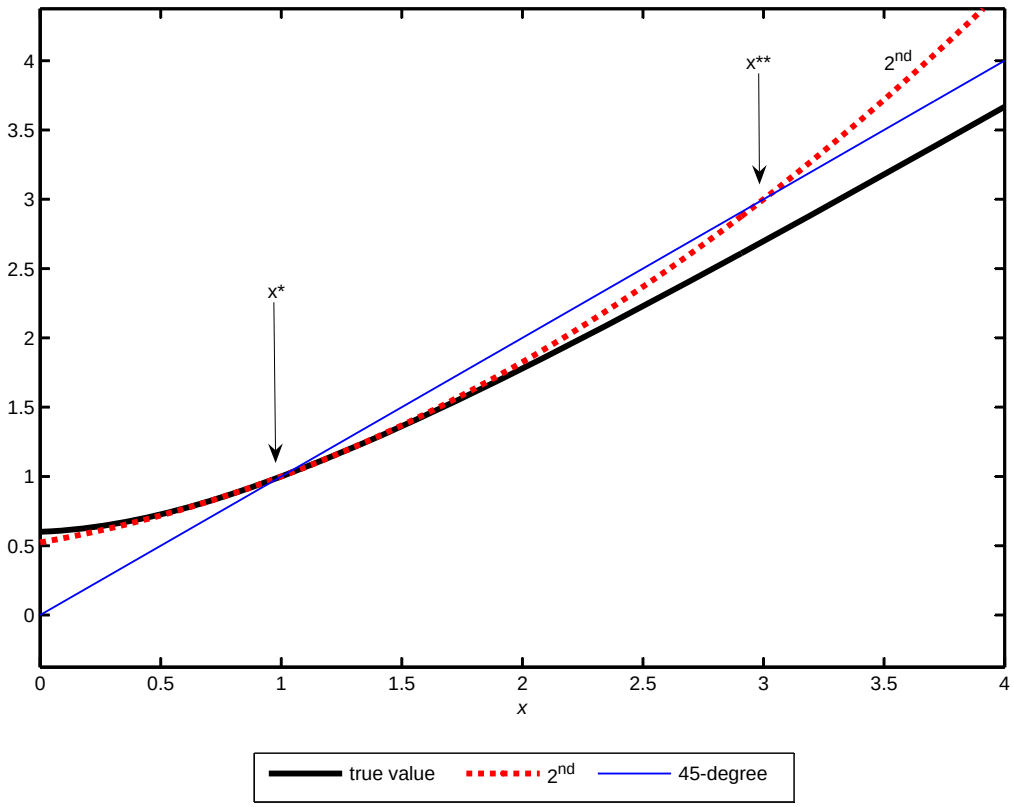
Notes: The graphs plot $x^{\alpha} + x$ for α equal to 5 and 11 and the Taylor series expansions at the indicated order.

Figure 2.B: $x^5 + x$ and 2^{nd}-order projections polynomial approximations



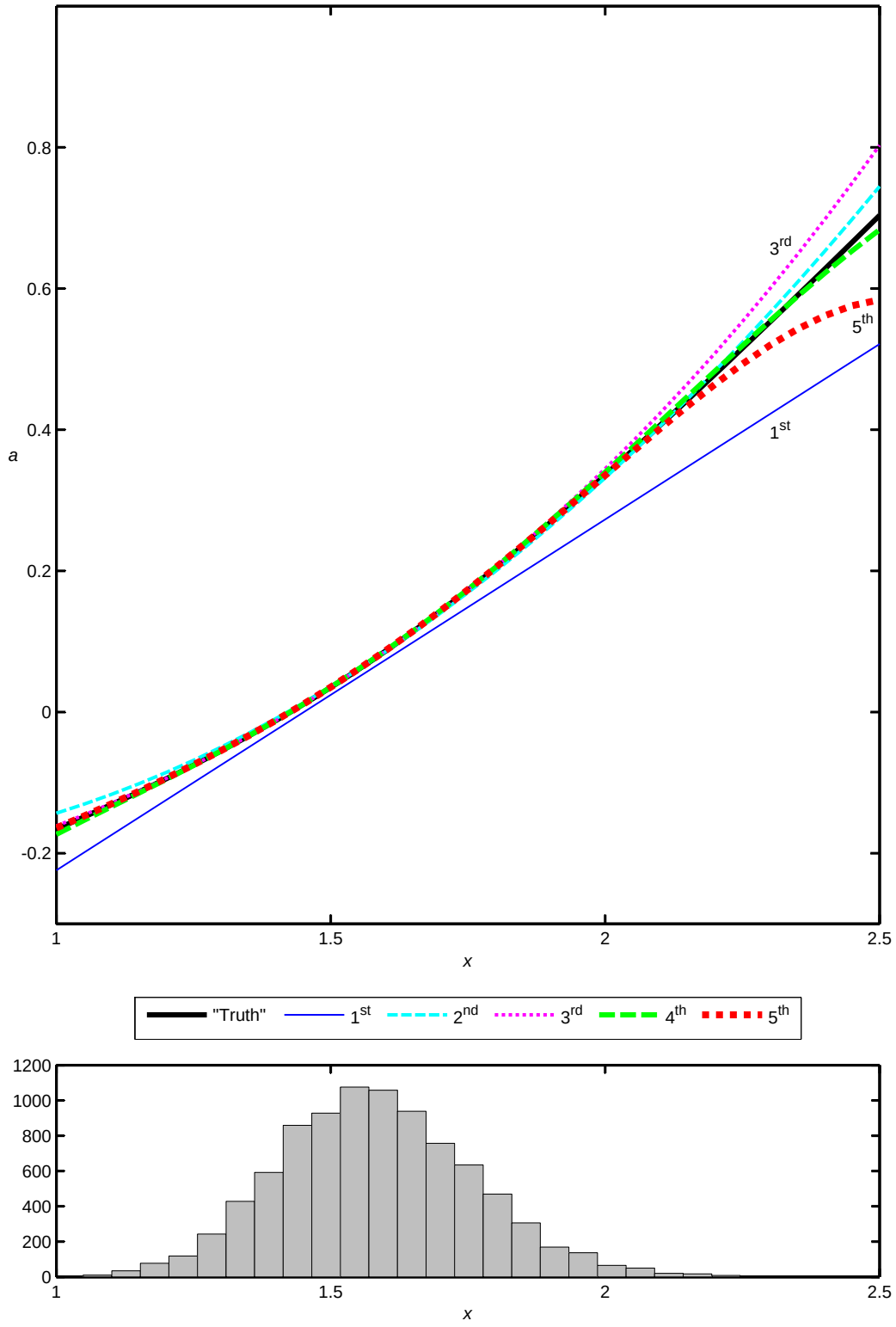
Notes: The graphs plot x^5+x and and the approximation by fitting a 2^{nd}-order polynomial using the indicated number of Chebyshev nodes.

Figure 3: $\alpha_0 + x + \alpha_1 \exp(-\alpha_2 x)$ and its 2^{nd}-order Taylor series approximation



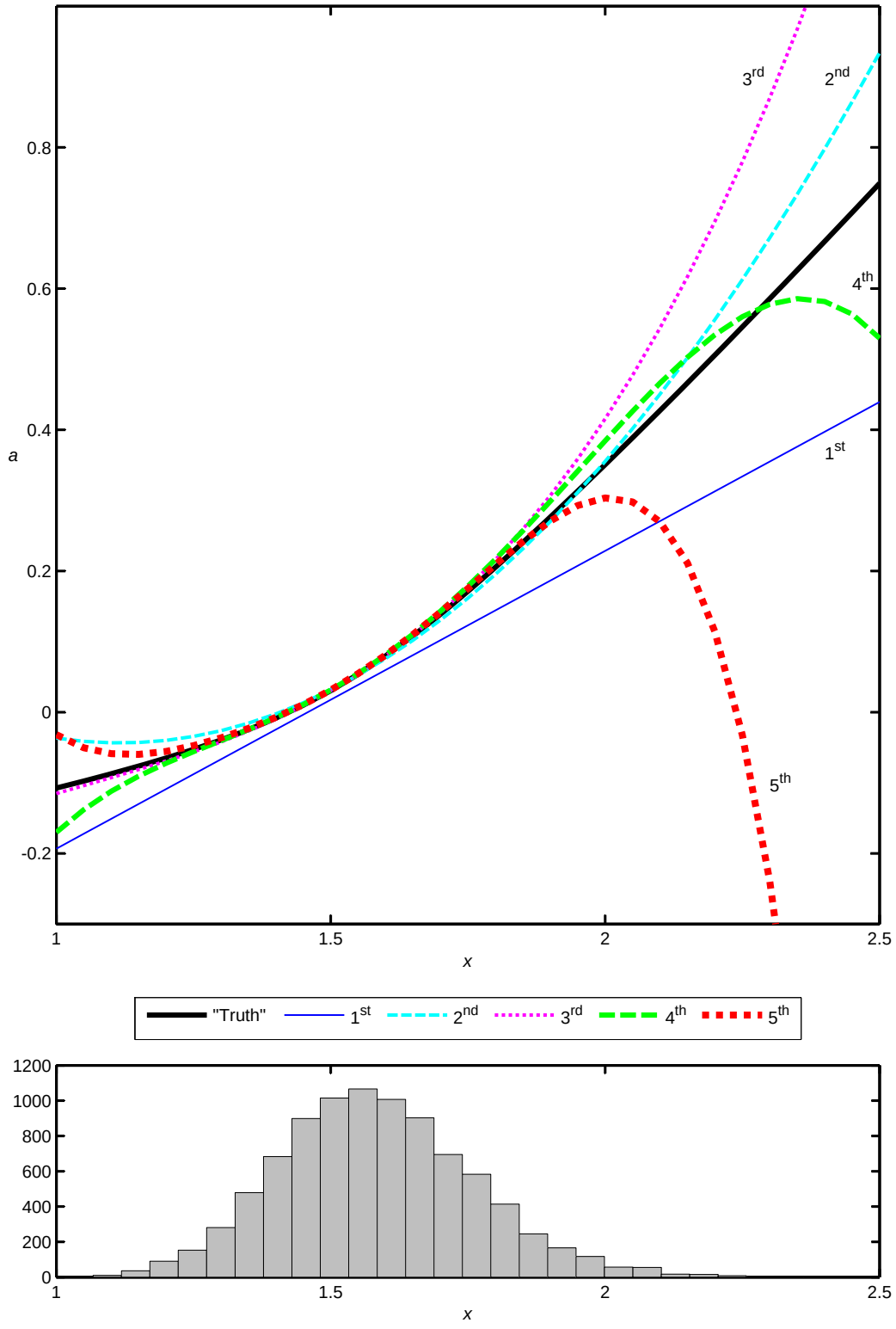
Notes: The values of α_0, α_1, and α_2 are equal to -0.3495, 0.95, and 1.

Figure 4.A: Perturbation solutions; $\eta = 10$



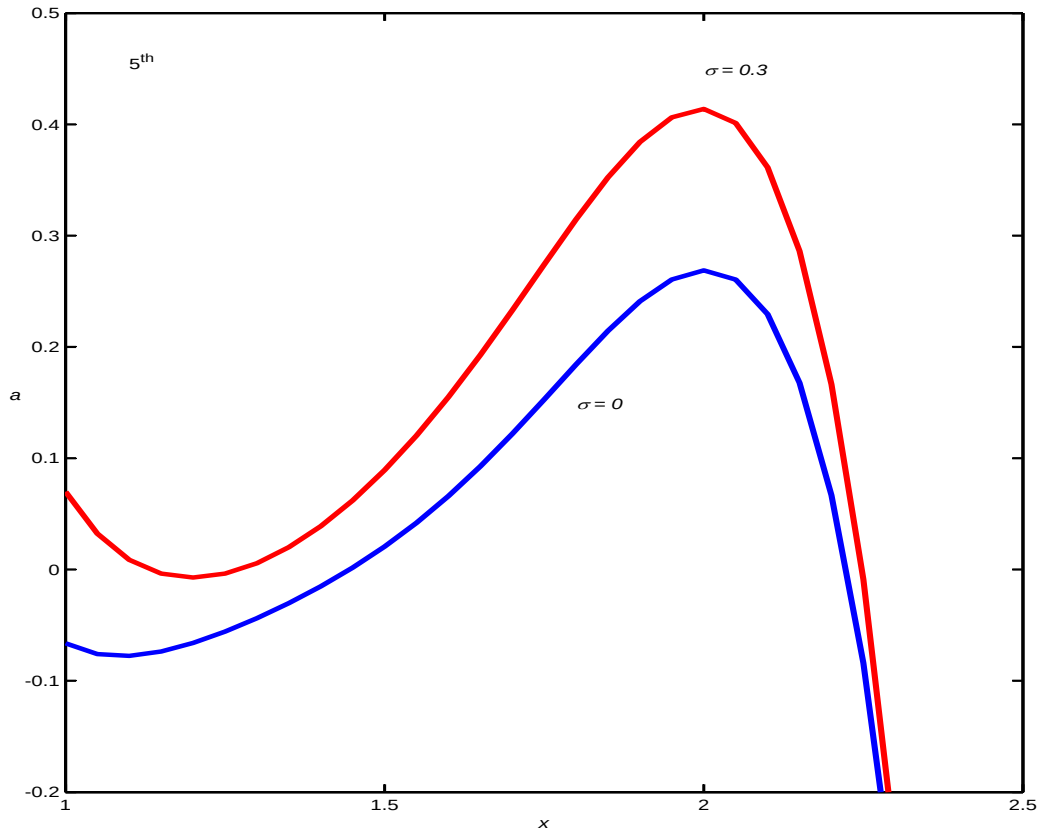
Notes: The top panel reports the perturbation solutions and the bottom panel reports the histogram of the state variable obtained using a sample of 9,000 observations generated with a very accurate solution.

Figure 4.B: Perturbation solutions; $\eta = 20$



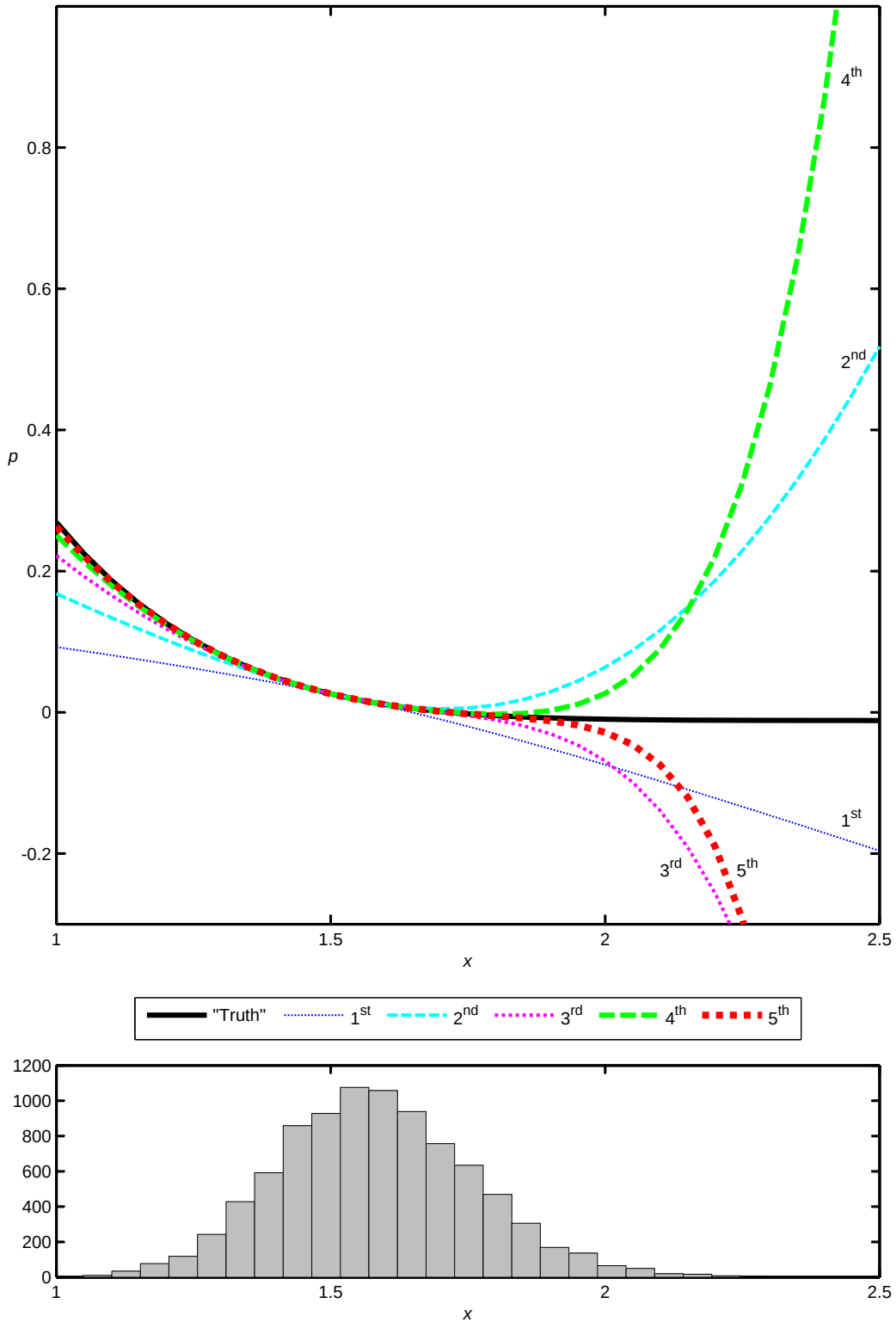
Notes: The top panel reports the perturbation solutions and the bottom panel reports the histogram of the state variable obtained using a sample of 9,000 observations generated with a very accurate solution.

Figure 5: Perturbation solutions for different levels of uncertainty; $\eta = 20$



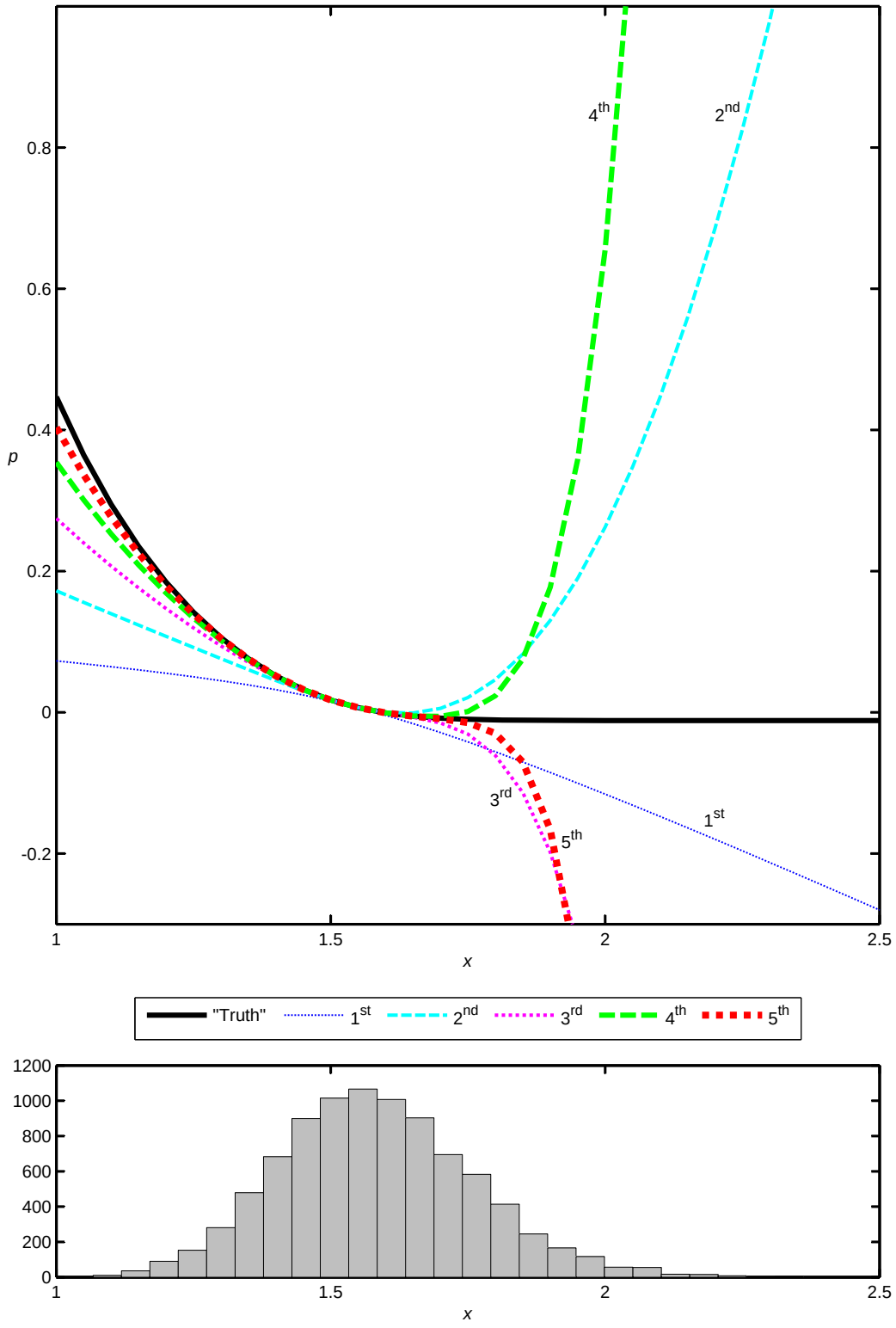
Notes: This graph plots the 5^{th}-order perturbation solution for the indicated standard deviation.

Figure 6.A: Taylor series approximation of $\partial P(a)/\partial a$; $\eta = 10$



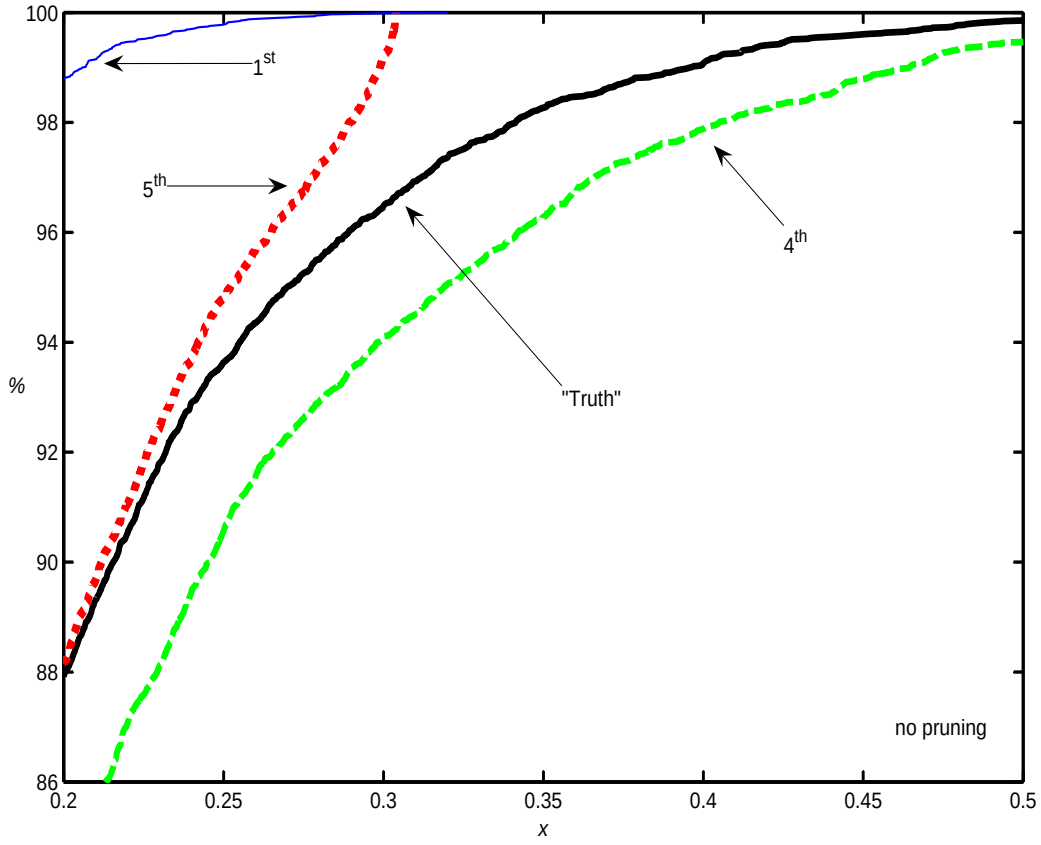
Notes: To make this graph comparable with Figure 4.A, we plot $\partial P(a)/\partial a$ evaluated at $a = h(x)$, where $h(x)$ is a very accurately solved policy rule. Since the same $h(x)$ is used for all perturbation solutions and, as indicated in Figure 4.A, $h(x)$ is a simple monotone increasing function, the graph still highlights the ability of using Taylor series expansions to approximate the penalty term. The bottom panel reports the histogram of the state variable obtained using a sample of 9,000 observations generated with a very accurate solution.

Figure 6.B: Taylor series approximation of $\partial P(a)/\partial a$; $\eta = 20$



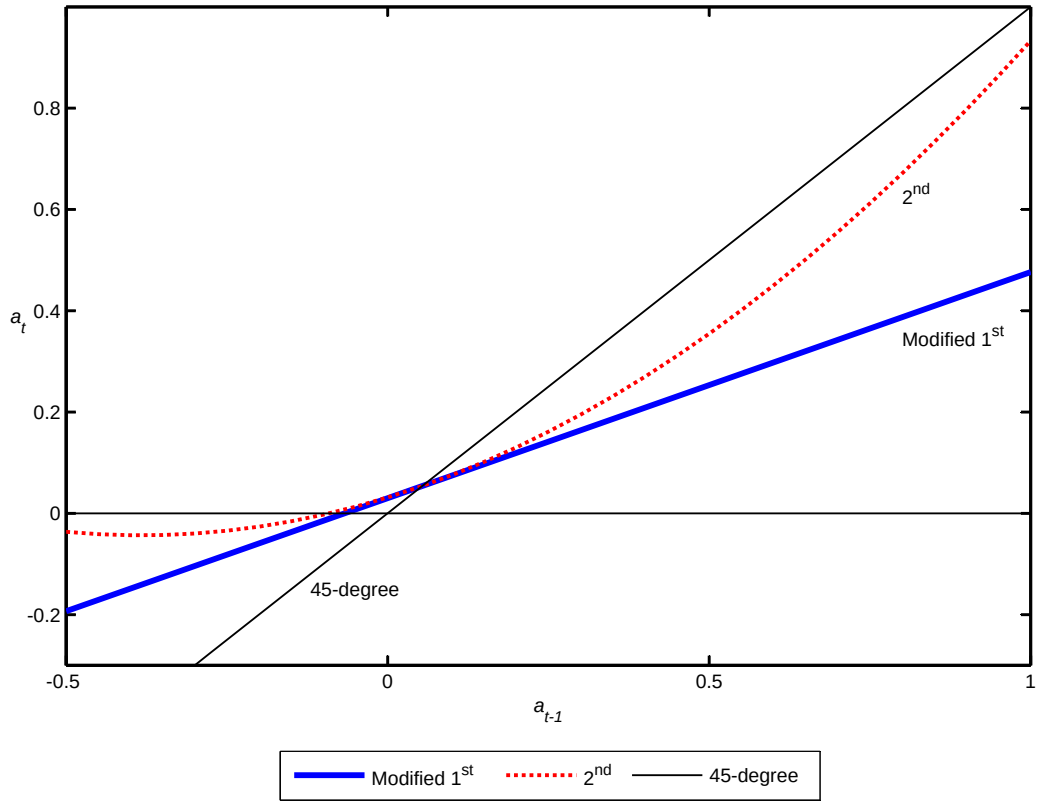
Notes: To make this graph comparable with Figure 4.B, we plot $\partial P(a)/\partial a$ evaluated at $a = h(x)$, where $h(x)$ is a very accurately solved policy rule. Since the same $h(x)$ is used for all perturbation solutions and, as indicated in Figure 4.B, $h(x)$ is a simple monotone increasing function, the graph still highlights the ability of using Taylor series expansions to approximate the penalty term. The bottom panel reports the histogram of the state variable obtained using a sample of 9,000 observations generated with a very accurate solution.

Figure 7: CDF of assset holdings for different numerical solutions; $\eta = 20$



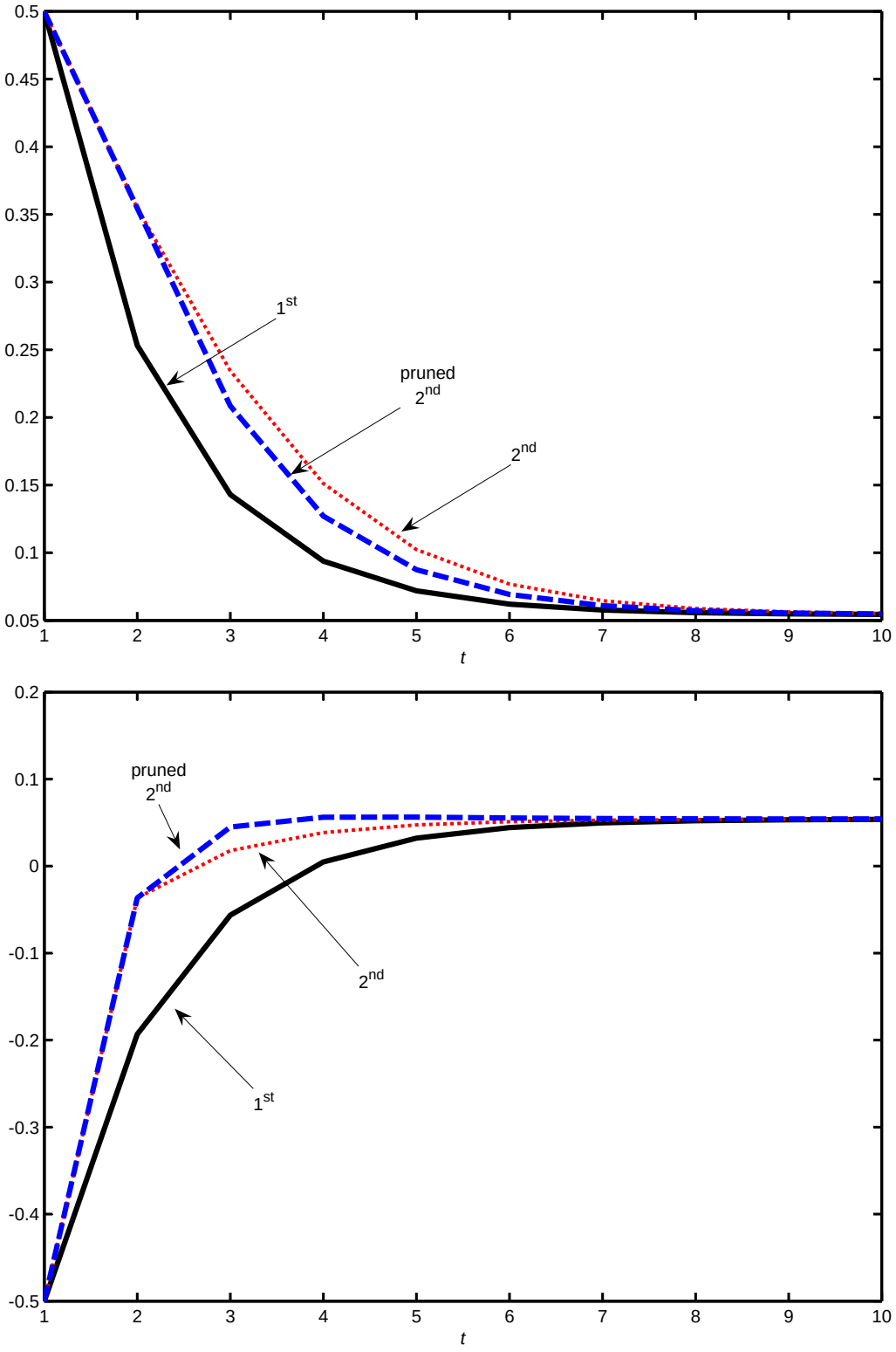
Notes: The CDFs are obtained using a simulation of 9,000 observations. For $\eta = 20$, the data simulated with the not-pruned 2^{nd} and 3^{rd}-order policy functions exploded and the CDF could not be calculated.

Figure 8: 1st and 2nd-order policy function as a function of a_{t-1}; $\eta = 20$



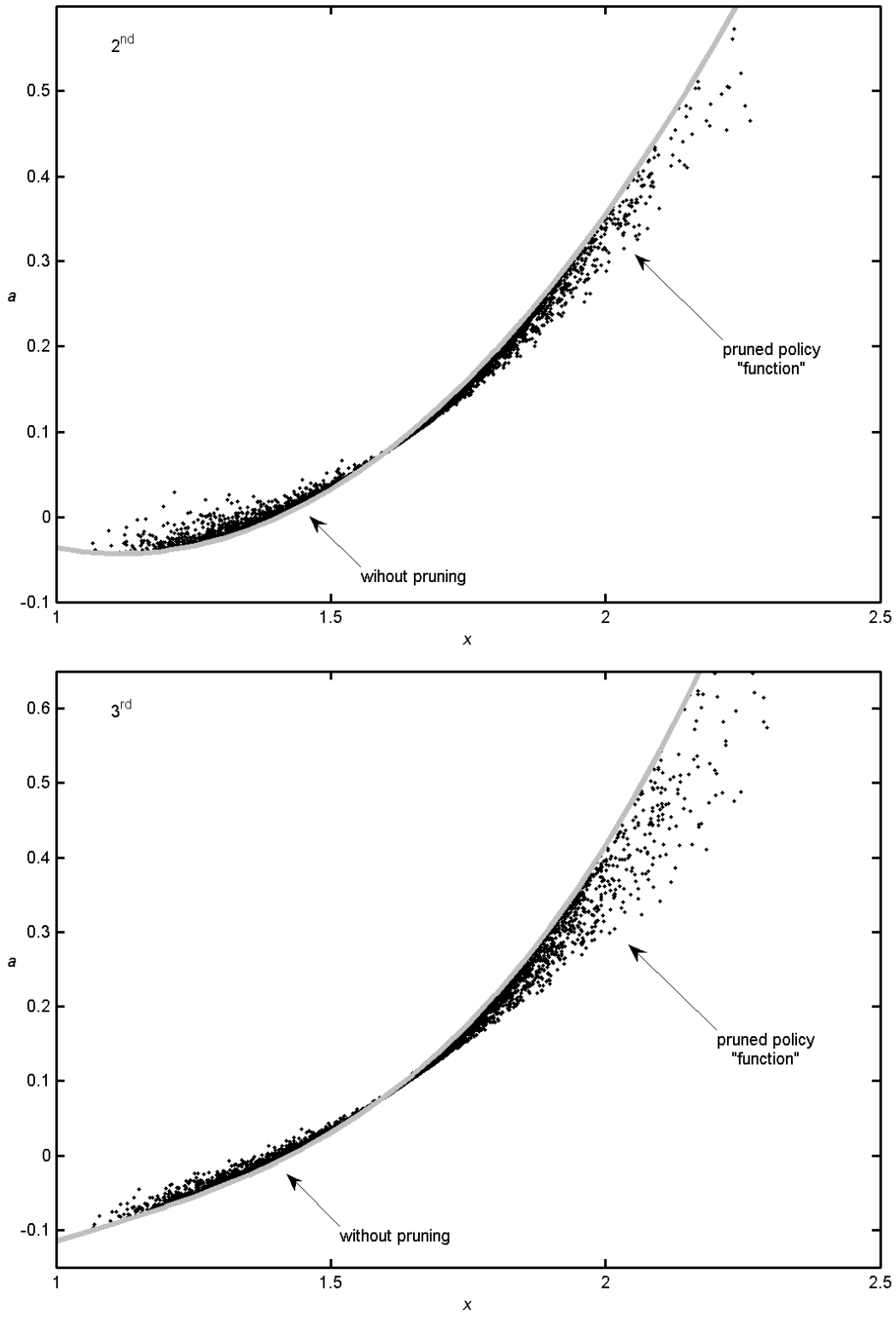
Notes: This graph plots the two Taylor series approximation as a function of beginning-of-period savings, a_{t-1}, instead of cash-on-hand, x, setting income equal to it's steady state value

Figure 9: Convergence towards steady state starting at $x = 0.5$ and $x = -0.5$



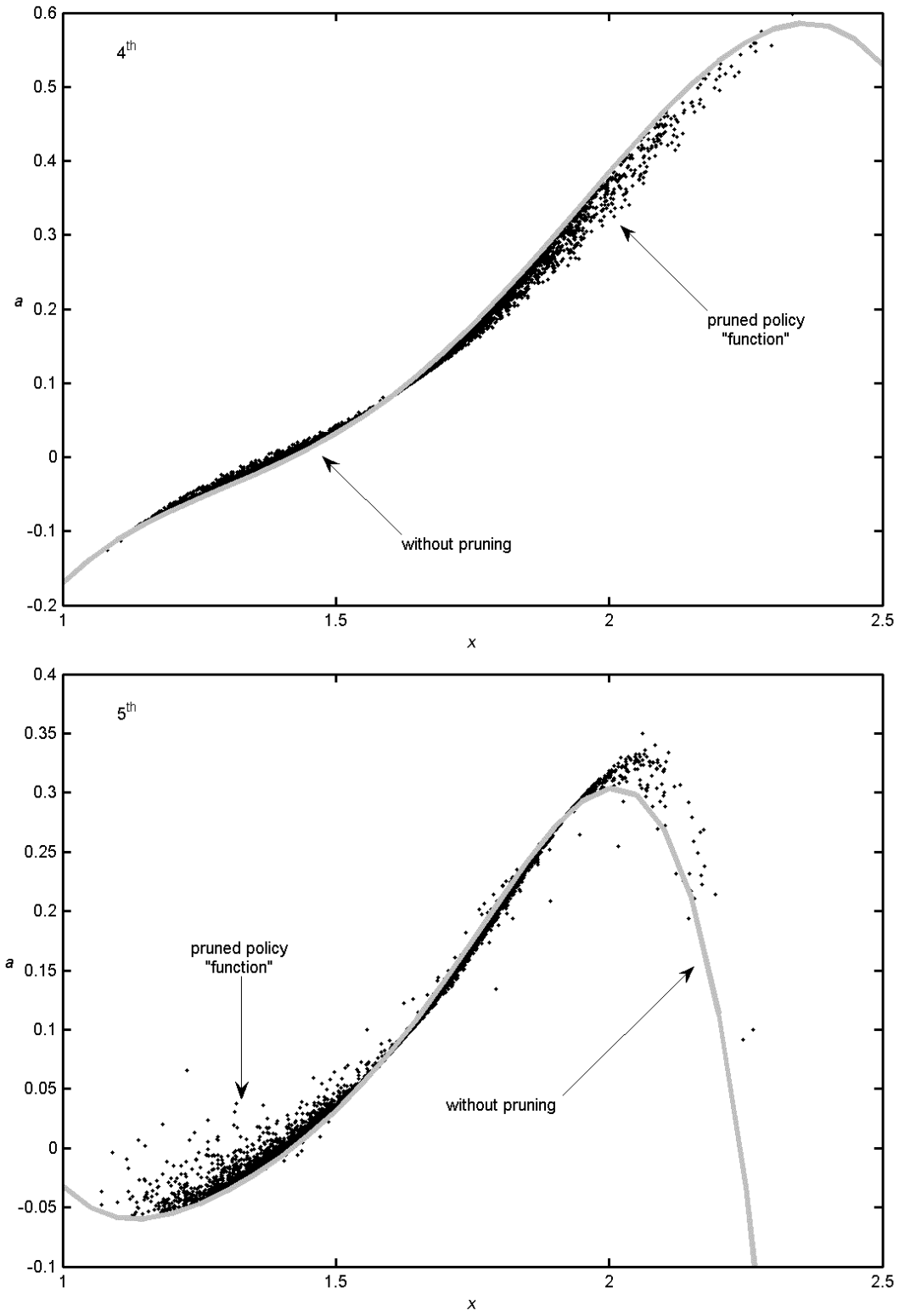
Notes: These graphs plot the time series according to the indicated numerical solution when the income level is kept fixed at its steady state level.

Figure 10.A: Implied policy "function" when pruning is used; $\eta = 20$



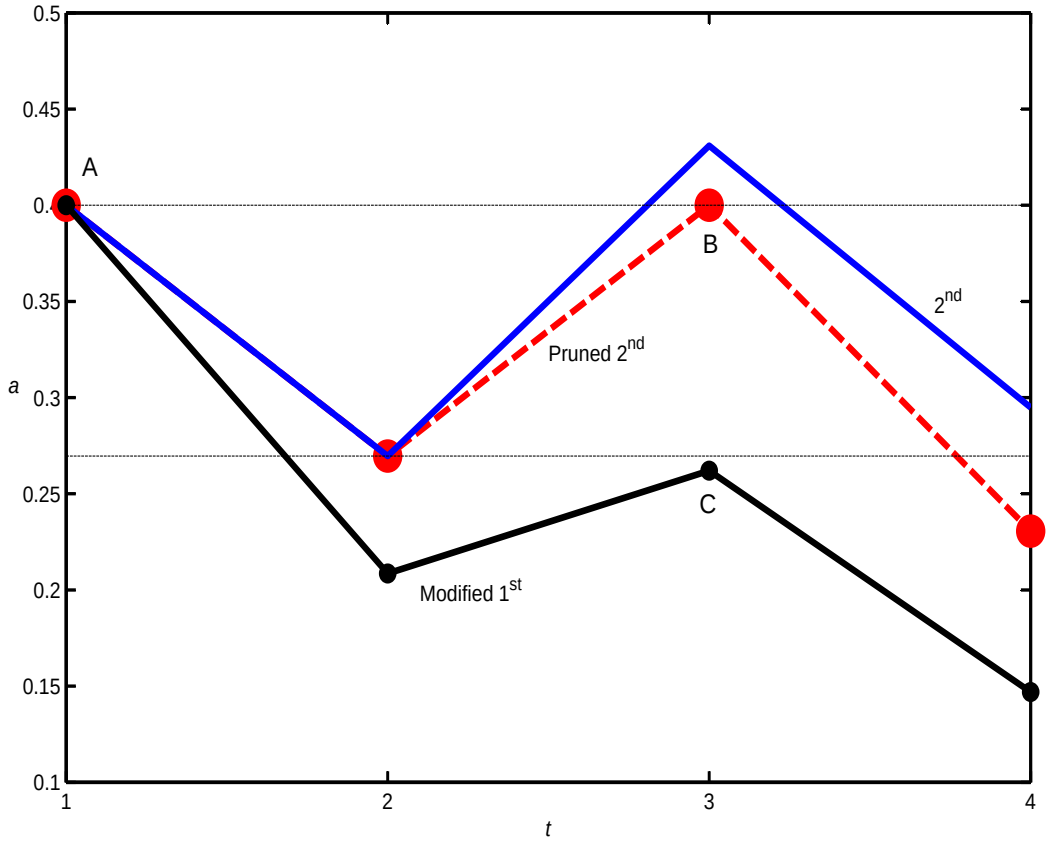
Notes: This graph plots the asset holdings chosen with the pruning procedure as a function of the observed state variable. It also plots the value corresponding to the perturbation solution underlying the pruning procedure.

Figure 10.B: Implied policy "function" when pruning is used; $\eta = 20$



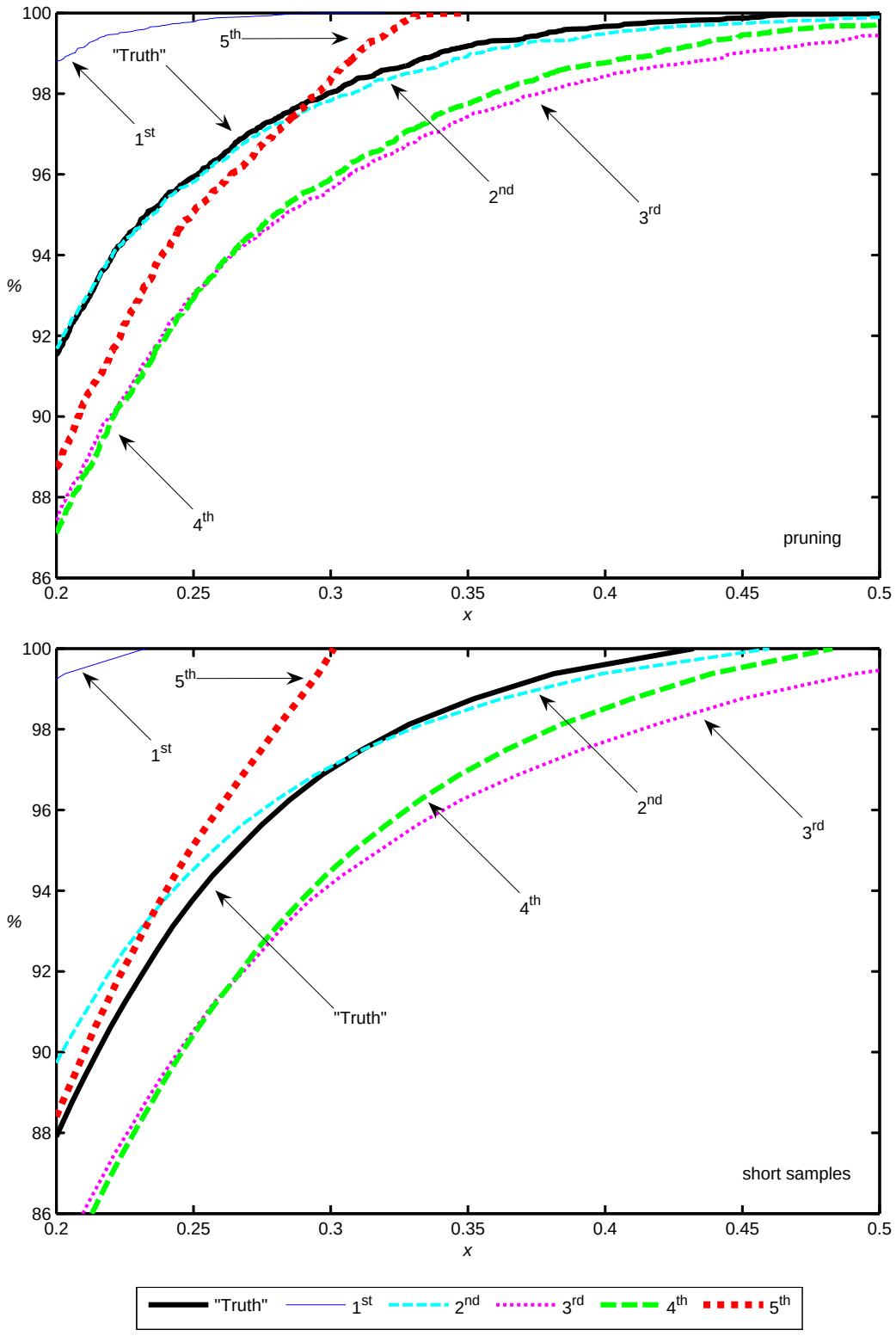
Notes: This graph plots the asset holdings chosen with the pruning procedure as a function of the observed state variable. It also plots the value corresponding to the perturbation solution underlying the pruning procedure.

Figure 11: Explaining why the pruning policy rule is not a function



Notes: This graph plots for $\eta = 20$ the time path for the chosen level of a according to the 1^{st}, 2^{nd}, and pruned 2^{nd}-order perturbation solution. θ is equal to its steady state value for $t = 2$ and a_t decreases towards its steady state value. For $t = 3$, the value of θ is such that for the pruned 2^{nd}-order solution the state variable in period 4, a_3 is equal to the state variable in period 2, a_1.

Figure 12: CDF of assset holdings for different numerical solutions; $\eta = 20$



Notes: The CDFs are obtained using a simulation of 9,000 observations. For $\eta = 20$, the data simulated with the not-pruned 2^{nd} and 3^{rd}-order policy functions exploded and the CDF could not be calculated.

Table 1: model properties - obtained using one long sample

	"Truth"	**Pertubation**				
		1st	**2**nd	**3**rd	**4**th	**5**th
$\eta_0 = \mathbf{10}$						
μ_a	0.0849	0.0472	0.0841	0.0868	0.0864	0.0857
σ_a	0.0973	0.0860	0.0934	0.1003	0.0985	0.0969
$a_{\min}$	−0.1491	−0.2180	−0.1289	−0.1461	−0.1539	−0.1475
$a_{\max}$	0.5900	0.4074	0.6000	0.6561	0.5921	0.5431
$\rho\left(a_{t+1}, y_t\right)$	0.8319	0.8677	0.8396	0.8202	0.8283	0.8336
$\rho\left(c_t, y_t\right)$	0.8373	0.8677	0.8470	0.8302	0.8342	0.8380
$\rho\left(c_t, a_{t+1}\right)$	0.9778	1.0000	0.9731	0.9681	0.9766	0.9798
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.8994	0.9190	0.9031	0.8940	0.8977	0.9002
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8944	0.9190	0.8984	0.8877	0.8923	0.8953
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.9675	1.0000	0.9616	0.9551	0.9658	0.9698
$\eta_0 = \mathbf{20}$						
μ_a	0.0849	0.0295	–	–	0.0965	0.0826
σ_a	0.0973	0.0698	–	–	0.1112	0.0869
$a_{\min}$	−0.0941	−0.1844	–	–	−0.1288	−0.4951
$a_{\max}$	0.6352	0.3202	–	–	0.5796	0.3038
$\rho\left(a_{t+1}, y_t\right)$	0.8127	0.9067	–	–	0.7676	0.8353
$\rho\left(c_t, y_t\right)$	0.8358	0.9067	–	–	0.8031	0.8476
$\rho\left(c_t, a_{t+1}\right)$	0.9281	1.0000	–	–	0.8985	0.8951
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.8873	0.9406	–	–	0.8673	0.8846
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8759	0.9406	–	–	0.8456	0.8579
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.8950	1.0000	–	–	0.8518	0.8109

Notes: These tables report summary statistics based on a sample of 9,000 observations. With "truth" we mean the policy rule obtained with a very accurate projections solution. μ_a is the mean of a, σ_a is the standard deviation, and $\rho(x_t, y_t)$ is the correlation between x_t and y_t.

Table 2: model properties - obtained using pruning and one long sample

	"Truth"	**Pertubation**				
		1st	**2**nd	**3**rd	**4**th	**5**th
$\eta_0 = \mathbf{10}$						
μ_a	0.0820	0.0472	0.0836	0.0858	0.0858	0.0854
σ_a	0.0926	0.0860	0.0910	0.0970	0.0962	0.0950
$a_{\min}$	−0.1407	−0.2180	−0.1230	−0.1373	−0.1476	−0.1393
$a_{\max}$	0.5632	0.4074	0.5827	0.6308	0.5821	0.5439
$\rho\left(a_{t+1}, y_t\right)$	0.8484	0.8677	0.8502	0.8351	0.8382	0.8414
$\rho\left(c_t, y_t\right)$	0.8451	0.8677	0.8491	0.8330	0.8362	0.8397
$\rho\left(c_t, a_{t+1}\right)$	0.9801	1.0000	0.9748	0.9725	0.9781	0.9797
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.9064	0.9190	0.9073	0.8995	0.9015	0.9034
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8968	0.9190	0.8977	0.8870	0.8915	0.8944
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.9687	1.0000	0.9625	0.9577	0.9665	0.9665
$\eta_0 = \mathbf{20}$						
μ_a	0.0752	0.0295	0.0796	0.0890	0.0901	0.0848
σ_a	0.0844	0.0698	0.0821	0.1036	0.0991	0.0838
$a_{\min}$	−0.0883	−0.1844	−0.0426	−0.0973	−0.1266	−0.1465
$a_{\max}$	0.5630	0.3202	0.6465	0.8863	0.6205	0.3500
$\rho\left(a_{t+1}, y_t\right)$	0.8607	0.9067	0.8478	0.8002	0.8182	0.8489
$\rho\left(c_t, y_t\right)$	0.8546	0.9067	0.8588	0.7906	0.8143	0.8559
$\rho\left(c_t, a_{t+1}\right)$	0.9366	1.0000	0.8947	0.8681	0.9138	0.9097
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.9058	0.9406	0.8966	0.8698	0.8845	0.8987
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8813	0.9406	0.8752	0.8054	0.8458	0.8755
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.8966	1.0000	0.8468	0.7801	0.8604	0.8622

Notes: These tables report summary statistics based on a sample of 9,000 observations. With "truth" we mean the policy rule obtained with a very accurate projections solution. μ_a is the mean of a, σ_a is the standard deviation, and $\rho(x_t, y_t)$ is the correlation between x_t and y_t.

Table 3: model properties - obtained using short samples - $\Lambda = 2$

	"Truth"	**Pertubation**				
		1st	**2**nd	**3**rd	**4**th	**5**th
$\eta_0 = \mathbf{10}$						
μ_a	0.0851	0.0473	0.0843	0.0870	0.0866	0.0859
	(0.0141)	(0.0116)	(0.0133)	(0.0149)	(0.0144)	(0.0140)
σ_a	0.0963	0.0852	0.0925	0.0991	0.0975	0.0960
	(0.0092)	(0.0060)	(0.0093)	(0.0109)	(0.0095)	(0.0088)
$a_{\min}$	-0.1066	-0.1580	-0.0944	-0.1056	-0.1081	-0.1055
	(0.0209)	(0.0294)	(0.0175)	(0.0201)	(0.0222)	(0.0207)
$a_{\max}$	0.4064	0.2939	0.4029	0.4319	0.4127	0.3987
	(0.0738)	(0.0464)	(0.0789)	(0.0909)	(0.0737)	(0.0627)
$\rho\left(a_{t+1}, y_t\right)$	0.8350	0.8686	0.8433	0.8254	0.8317	0.8363
	(0.0121)	(0.0022)	(0.0139)	(0.0181)	(0.0127)	(0.0103)
$\rho\left(c_t, y_t\right)$	0.8373	0.8686	0.8468	0.8299	0.8343	0.8383
	(0.0099)	(0.0022)	(0.0109)	(0.0134)	(0.0103)	(0.0088)
$\rho\left(c_t, a_{t+1}\right)$	0.9802	1.0000	0.9762	0.9721	0.9792	0.9819
	(0.0040)	(0.0000)	(0.0064)	(0.0096)	(0.0041)	(0.0027)
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.9009	0.9197	0.9047	0.8958	0.8992	0.9016
	(0.0113)	(0.0070)	(0.0117)	(0.0136)	(0.0116)	(0.0106)
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8941	0.9197	0.8978	0.8869	0.8920	0.8953
	(0.0110)	(0.0070)	(0.0118)	(0.0144)	(0.0114)	(0.0102)
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.9692	1.0000	0.9634	0.9572	0.9675	0.9715
	(0.0067)	(0.0000)	(0.0101)	(0.0143)	(0.0068)	(0.0051)
Rejection	0%	-	0%	0%	0%	0%
$\eta_0 = \mathbf{20}$						
μ_a	0.0835	0.0297	0.0814	0.0853	0.0888	0.0827
	(0.0125)	(0.0086)	(0.0118)	(0.0107)	(0.0121)	(0.0119)
σ_a	0.0949	0.0693	0.0892	0.1034	0.1024	0.0867
	(0.0101)	(0.0046)	(0.0123)	(0.0105)	(0.0099)	(0.0062)
$a_{\min}$	-0.0706	-0.1368	-0.0401	-0.0742	-0.0836	-0.0639
	(0.0121)	(0.0231)	(0.0033)	(0.0117)	(0.0213)	(0.0527)
$a_{\max}$	0.4250	0.2325	0.4427	0.4893	0.4601	0.3003
	(0.0785)	(0.0380)	(0.1028)	(0.0887)	(0.0691)	(0.0063)
$\rho\left(a_{t+1}, y_t\right)$	0.8216	0.9073	0.8181	0.7630	0.7886	0.8345
	(0.0194)	(0.0015)	(0.0357)	(0.0383)	(0.0234)	(0.0423)
$\rho\left(c_t, y_t\right)$	0.8365	0.9073	0.8492	0.7969	0.8093	0.8508
	(0.0150)	(0.0015)	(0.0209)	(0.0221)	(0.0188)	(0.0190)
$\rho\left(c_t, a_{t+1}\right)$	0.9351	1.0000	0.8931	0.8617	0.9085	0.9052
	(0.0094)	(0.0000)	(0.0284)	(0.0367)	(0.0112)	(0.1120)
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.8909	0.9411	0.8857	0.8602	0.8756	0.8853
	(0.0143)	(0.0051)	(0.0202)	(0.0208)	(0.0161)	(0.0538)
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8764	0.9411	0.8737	0.8301	0.8508	0.8728
	(0.0150)	(0.0051)	(0.0227)	(0.0251)	(0.0200)	(0.0412)
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.9008	1.0000	0.8482	0.8022	0.8632	0.8446
	(0.0163)	(0.0000)	(0.0361)	(0.0410)	(0.0218)	(0.1634)
Rejection	7.8%	-	9.9%	94.8%	43.3%	1.0%

Notes: Standard deviation across MC replications in parentheses. Λ is the parameter of the rejection criterion. For further information see notes of Table 1.

Table 4: model properties -obtained using short samples - $\Lambda = 3$

	"Truth"	**Pertubation**				
		1st	**2**nd	**3**rd	**4**th	**5**th
$\eta_0 = \mathbf{10}$						
μ_a	0.0851	0.0473	0.0843	0.0870	0.0866	0.0859
	(0.0141)	(0.0116)	(0.0133)	(0.0149)	(0.0144)	(0.0140)
σ_a	0.0963	0.0852	0.0925	0.0991	0.0975	0.0960
	(0.0092)	(0.0060)	(0.0093)	(0.0109)	(0.0095)	(0.0088)
$a_{\min}$	-0.1066	-0.1580	-0.0944	-0.1056	-0.1081	-0.1055
	(0.0209)	(0.0294)	(0.0175)	(0.0201)	(0.0222)	(0.0207)
$a_{\max}$	0.4064	0.2939	0.4029	0.4319	0.4127	0.3987
	(0.0738)	(0.0464)	(0.0789)	(0.0909)	(0.0737)	(0.0627)
$\rho\left(a_{t+1}, y_t\right)$	0.8350	0.8686	0.8433	0.8254	0.8317	0.8363
	(0.0121)	(0.0022)	(0.0139)	(0.0181)	(0.0127)	(0.0103)
$\rho\left(c_t, y_t\right)$	0.8373	0.8686	0.8468	0.8299	0.8343	0.8383
	(0.0099)	(0.0022)	(0.0109)	(0.0134)	(0.0103)	(0.0088)
$\rho\left(c_t, a_{t+1}\right)$	0.9802	1.0000	0.9762	0.9721	0.9792	0.9819
	(0.0040)	(0.0000)	(0.0064)	(0.0096)	(0.0041)	(0.0027)
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.9009	0.9197	0.9047	0.8958	0.8992	0.9016
	(0.0113)	(0.0070)	(0.0117)	(0.0136)	(0.0116)	(0.0106)
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8941	0.9197	0.8978	0.8869	0.8920	0.8953
	(0.0110)	(0.0070)	(0.0118)	(0.0144)	(0.0114)	(0.0102)
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.9692	1.0000	0.9634	0.9572	0.9675	0.9715
	(0.0067)	(0.0000)	(0.0101)	(0.0143)	(0.0068)	(0.0051)
Rejection	0%	-	0%	0%	0%	0%
$\eta_0 = \mathbf{20}$						
μ_a	0.0851	0.0297	0.0834	0.0957	0.0968	0.0827
	(0.0144)	(0.0086)	(0.0141)	(0.0191)	(0.0179)	(0.0119)
σ_a	0.0962	0.0693	0.0920	0.1155	0.1093	0.0868
	(0.0116)	(0.0046)	(0.0163)	(0.0219)	(0.0146)	(0.0068)
$a_{\min}$	-0.0705	-0.1368	-0.0401	-0.0756	-0.0823	-0.0649
	(0.0120)	(0.0231)	(0.0033)	(0.0131)	(0.0216)	(0.0616)
$a_{\max}$	0.4323	0.2325	0.4598	0.5558	0.4828	0.3011
	(0.0849)	(0.0380)	(0.1251)	(0.1451)	(0.0720)	(0.0258)
$\rho\left(a_{t+1}, y_t\right)$	0.8196	0.9073	0.8111	0.7235	0.7777	0.8339
	(0.0210)	(0.0015)	(0.0459)	(0.0764)	(0.0275)	(0.0462)
$\rho\left(c_t, y_t\right)$	0.8351	0.9073	0.8462	0.7757	0.8023	0.8503
	(0.0162)	(0.0015)	(0.0244)	(0.0401)	(0.0219)	(0.0252)
$\rho\left(c_t, a_{t+1}\right)$	0.9348	1.0000	0.8881	0.8180	0.9069	0.9036
	(0.0096)	(0.0000)	(0.0361)	(0.0858)	(0.0110)	(0.1233)
$\rho\left(\Delta a_{t+1}, \Delta y_t\right)$	0.8899	0.9411	0.8827	0.8488	0.8700	0.8845
	(0.0151)	(0.0051)	(0.0243)	(0.0324)	(0.0183)	(0.0595)
$\rho\left(\Delta c_t, \Delta y_t\right)$	0.8753	0.9411	0.8709	0.8083	0.8449	0.8721
	(0.0159)	(0.0051)	(0.0261)	(0.0447)	(0.0227)	(0.0459)
$\rho\left(\Delta c_t, \Delta a_{t+1}\right)$	0.8999	1.0000	0.8429	0.7669	0.8570	0.8429
	(0.0169)	(0.0000)	(0.0429)	(0.0761)	(0.0245)	(0.1725)
Rejection	0%	-	0.6%	81%	0.1%	0.9%

Notes: Standard deviation across MC replications in parentheses. Λ is the parameter of the rejection criterion. For further information see notes of Table 1.

Table 5: Accuracy test - Euler equation errors

	"Truth"	**Pertubation**				
		1st	**2**nd	**3**rd	**4**th	**5**th
$\eta_0 = \mathbf{10}$						
Maximum error	$5.11E-6$	0.1614	0.0521	0.0403	0.0188	0.0832
Average error	$1.64E-6$	0.0456	0.0090	0.0079	0.0027	0.0084
$\eta_0 = \mathbf{20}$						
Maximum error	$1.40E-5$	0.6250	0.2273	∞	0.4068	∞
Average error	$4.21E-6$	0.1113	0.0335	∞	0.0354	∞

Notes: This table reports the Euler equation errors expressed as consumption equivalents using a grid of 15,000 equidistant nodes on $[1, 2.5]$. If a procedure predicted consumption to be negative when calculating the conditional expectation and the marginal utility, thus could not be calcuatled, then we report this with ∞ nodes distributed

Previous DNB Working Papers in 2009

No. 198 **Peter ter Berg**, Unification of the Fréchet and Weibull Distribution
No. 199 **Ronald Heijmans**, Simulations in the Dutch interbank payment system: A sensitivity analysis
No 200 **Itai Agur**, What Institutional Structure for the Lender of Last Resort?
No 201 **Iman van Lelyveld, Franka Liedorp and Manuel Kampman**, An empirical assessment of reinsurance risk
No. 202 **Kerstin Bernoth and Andreas Pick**, Forecasting the fragility of the banking and insurance sector
No. 203 **Maria Demertzis**, The 'Wisdom of the Crowds' and Public Policy
No. 204 **Wouter den Haan and Vincent Sterk**, The comovement between household loans and real activity
No. 205 **Gus Garita and Chen Zhou**, Can Open Capital Markets Help Avoid Currency Crises?
No. 206 **Frederick van der Ploeg and Steven Poelhekke**, The Volatility Curse: Revisiting the Paradox of Plenty
No. 207 **M. Hashem Pesaran and Adreas Pick**, Forecasting Random Walks under Drift Instability
No. 208 **Zsolt Darvas**, Monetary Transmission in three Central European Economies: Evidence from Time-Varying Coefficient Vector Autoregressions
No. 209 **Steven Poelhekke**, Human Capital and Employment Growth in German Metropolitan Areas: New Evidence
No. 210 **Vincent Sterk**, Credit Frictions and the Comovement between Durable and Non-durable Consumption
No. 211 **Jan de Dreu and Jacob Bikker**, Pension fund sophistication and investment policy
No. 212 **Jakob de Haan and David-Jan Jansen**, The communication policy of the European Central Bank: An overview of the first decade
No. 213 **Itai Agur**, Regulatory Competition and Bank Risk Taking
No. 214 **John Lewis**, Fiscal policy in Central and Eastern Europe with real time data: Cyclicality, inertia and the role of EU accession
No. 215 **Jacob Bikker**, An extended gravity model with substitution applied to international trade
No. 216 **Arie Kapteyn and Federica Teppa**, Subjective Measures of Risk Aversion, Fixed Costs, and Portfolio Choice
No. 217 **Mark Mink and Jochen Mierau**, Measuring Stock Market Contagion with an Application to the Sub-prime Crisis
No. 218 **Michael Biggs, Thomas Mayer and Andreas Pick**, Credit and economic recovery
No. 219 **Chen Zhou**, Dependence structure of risk factors and diversification effects
No. 220 **W.L. Heeringa and A.L. Bovenberg**, Stabilizing pay-as-you-go pension schemes in the face of rising longevity and falling fertility: an application to the Netherlands
No. 221 **Nicole Jonker and Anneke Kosse**, The impact of survey design on research outcomes: A case study of seven pilots measuring cash usage in the Netherlands
No. 222 **Gabriele Galati, Steven Poelhekke and Chen Zhou**, Did the crisis affect inflation expectations?
No. 223 **Jacob Bikker, Dirk Broeders, David Hollanders and Eduard Ponds**, Pension funds' asset allocation and participants age: a test of the life-cycle model
No. 224 **Stijn Claessens and Neeltje van Horen,** Being a Foreigner among Domestic Banks: Asset or Liability?
No. 225 **Jacob Bikker, Sherrill Shaffer and Laura Spierdijk**, Assessing Competition with the Panzar-Rosse Model: The Role of Scale, Costs, and Equilibrium
No. 226 **Jan Marc Berk and Beata K. Bierut**, Communication in a monetary policy committee: a note
No. 227 **Dirk Broeders, An Chen and Birgit Koos**, An institutional evaluation of pension funds and life insurance companies
No. 228 **Michael Hurd, Maarten van Rooij and Joachim Winter**, Stock market expectations of Dutch households
No. 229 **Maria Demertzis, Massimiliano Marcellino and Nicola Viegi**, Anchors for Inflation Expectations
No. 230 **Jan Willem van den End and Mostafa Tabbae**, When liquidity risk becomes a macro-prudential issue: Empirical evidence of bank behaviour
No. 231 **Maarten van Rooij, Annamaria Lusardi and Rob Alessi,** Financial Literacy and Retirement Planning in the Netherlands
No.232 **Chen Zhou,** Are banks too big to fail?
No.233 **Frederick van der Ploeg, Steven Poelhekke**, THE PUNGENT SMELL OF "RED HERRINGS": Subsoil assets, rents, volatility and the resource curse
No.234 **Mark Mink,** Is Contagion in the Eye of the Beholder?
No.235 **Carin van der Cruijsen and Maria Demertzis** How Anchored Are Inflation Expectations in EMU Countries?
No.236 **Christian Huurman, Francesco Ravazzolo and Chen Zhou** The power of weather

No. 237 **Itai Agur, On** the Possibility of Credit Rationing in the Stiglitz-Weiss Model: A Comment
No. 238 **Janko Gorter en Paul Schilp,** Determinants of consumer financial risktaking: Evidence from deductible choice
No. 239 **Wilco Bolt and Heiko Schmiedel,** SEPA, Efficiency, and Payment Card Competition