Lower Bounds on Approximation Errors: Testing the Hypothesis That a Numerical Solution Is Accurate *

Kenneth L. Judd, Lilia Maliar and Serguei Maliar

August 20, 2014

Abstract

We propose a novel methodology for evaluating the accuracy of numerical solutions to dynamic economic models. Specifically, we construct a lower bound on the size of approximation errors. A small lower bound on errors is a necessary condition for accuracy: If a lower error bound is unacceptably large, then the actual approximation errors are even larger, and hence, we reject the hypothesis that a numerical solution is accurate. Our accuracy analysis is logically equivalent to hypothesis testing in statistics. As an illustration of our methodology, we assess approximation errors in the first- and second-order perturbation solutions for two stylized models: a neoclassical growth model and a new Keynesian model. The errors are small for the former model but unacceptably large for the latter model under some empirically relevant parameterizations.

JEL classification : C61, C63, C68, E31, E52

Key Words : approximation errors; best case scenario, error bounds, Euler equation residuals; accuracy; numerical solution; algorithm; new Keynesian model

*We benefited from comments of participants of the 2014 SITE workshop at Stanford University. Lilia Maliar and Serguei Maliar acknowledge support from the Hoover Institution and Department of Economics at Stanford University, University of Alicante, Ivie, Santa Clara University under the grant DPROV115, the Generalitat Valenciana under the grant Prometeo/2013/037 and the MECD under the grant ECO2012-36719. Corresponding author: Lilia Maliar, Office 249, Department of Economics, Stanford, CA 94305-6072, USA; email: maliarl@stanford.edu.

1 Introduction

Dynamic economic models do not typically admit closed-form solutions and must be studied with numerical methods. A numerical method approximates the true solution up to some degree of accuracy. The control over the quality of approximation is critical if we want to get valid inferences from numerical experiments. That is, the constructed approximation must have a minimum acceptable quality for the questions studied; otherwise, it could happen that our conclusions and policy implications are simply driven by approximation errors (or some bugs in the code).

In this paper, we propose a novel methodology for evaluating the accuracy of numerical solutions to dynamic economic models. Specifically, we construct a lower bound on the size of approximation errors in the model's variables (an approximation error is defined as a unit-free difference between a true solution and an approximation). A small lower bound on errors is a necessary condition for accuracy: If a lower error bound is unacceptably large, then the actual approximation errors are even larger, and hence, we conclude that a numerical solution is insufficiently accurate. Our methodology of error analysis is very general: it is independent of a specific solution method, and it is applicable to both dynamic programming and equilibrium problems.

Our accuracy testing is logically equivalent to hypothesis testing in statistics. We form a hypothesis that an approximate solution is accurate, and we test this hypothesis by assuming an optimistic – best-case – scenario for the approximation errors. If the approximation errors are unacceptable even under the best case scenario, we reject the hypothesis that an approximate solution is accurate. However, if a lower bound on errors is small and an approximate solution passes our accuracy check, we still cannot affirm that the solution is accurate. As in statistics, our accuracy testing is only meant to reject a hypothesis which is false but not to accept a hypothesis which is true.

As an illustration, we apply our methodology to assess the size of approximation errors in the first- and second-order perturbation solutions for two stylized models: a neoclassical optimal growth model and a new Keynesian model. For the growth model, we find that the approximation errors of the first-order perturbation solutions (linearization) are at most of order 0.1%, and they are even lower for the second-order perturbation solution. These errors are sufficiently small, and thus we cannot reject the hypothesis that perturbation methods are sufficiently accurate for the studied model example. However, for the new Keynesian models, the accuracy of the perturbation solutions depends on a specific parameterization used: the approximation errors are low if the volatility of shocks is low, however, they become unacceptably large when the volatility of shocks increases. The approximation errors can reach hundreds percent under some empirically relevant parameterizations. This finding shows that the accuracy implications obtained for

one class of models cannot be taken for granted for other classes of models and that accuracy evaluations are essential for every model studied. (In terms of hypothesis testing, we can make the following parallel statement: if regression coefficients are significant in one econometric model, it does not mean that they will significant in other econometric models).

There are three main approaches to accuracy evaluations in the literature. The first approach is a forward error analysis that poses the following question: Given an economic model, how much an approximate solution must be modified to satisfy all model's conditions exactly? A conventional forward error analysis constructs an upper bound on approximation errors by assuming the worst-case scenario. Upper error bounds are derived for policy function iteration (Bertsekas and Tsitsiklis (1996), p. 275), value-iterative methods (Santos and Vigo-Aguillar (1998), Santos (2000) and Santos and Peralta-Alva (2005)), for perturbation methods (Schmitt-Grohé and Uribe (2004)); see Santos and Peralta-Alva (2014) for a review. A small upper bound on errors is a sufficient condition for accuracy: If an upper error bound is small, then the actual approximation errors are even smaller, and one can accept the hypothesis that a numerical solution is accurate.

Our lower bound error analysis is also a forward error analysis but it provides a necessary rather than sufficient condition for accuracy. A potential shortcoming of the conventional upper bound error analysis is that worst-case scenario can be too pessimistic and may reject solutions that are sufficiently accurate. In turn, our best case scenario can be too optimistic and may fail to reject solutions that are insufficiently accurate. However, these features are not specific to the analysis of approximation errors but are common to all necessary and sufficient conditions in general. The upper- and lower-bound tests can be viewed as complementary. We shall also mention that lower error bounds are easy to construct, while upper error bounds are derived for specific models and specific numerical methods; and their extension to other applications is a non-trivial task.

The second approach to accuracy evaluation – a residual analysis – consists in a numerical evaluation of residuals in the model's equations such as first-order conditions, Euler equations, Bellman equation, constraints and laws of motions of exogenous shocks; see Judd (1992), Jin and Judd (2000), Aruoba et al. (2005), Juillard and Villemot (2011), Judd et al. (2011a), among others; and also, see a statistical residual test by Den Haan and Marcet (1994). The analysis of residuals is simple, general and inexpensive but has an important shortcoming: A relation between the size of the residuals and the size of approximation errors in model's variables is unknown with an exception of a special case of strongly concave infinite-horizon optimization problems studied in Santos (2000). In contrast, our lower error bounds are constructed for approximation errors in the model's variables – our true objects of interest.

Finally, the third approach to accuracy evaluation includes a backward error analysis and a mixed forward-backward error analysis introduced in Wilkinson (1963) and Higham (1996), respectively. A backward error analysis inverts the question posed by a forward error analysis: Given an approximate solution, how much an economic model itself (in terms of parameters) must be modified in order to make an approximate solution to satisfy all model's equations? A mixed forward-backward error analysis allows for modifications in equilibrium quantities, in addition to modifications in parameters. Kubler and Schmedders (2005) show how these two approaches can be applied to assess the difference between approximate and exact equilibria in a life-cycle model with incomplete markets and heterogeneous agents models. Sims (1990) proposes an accuracy test which is similar in spirit to the backward error analysis: he argues that the difference between approximate and exact equilibria can be measured by the difference between the distributions of the true stochastic shocks and the stochastic shocks that are implied by the approximate solution. Finally, a recent paper of Kogan and Mitra (2013) proposes a novel and promising technique of measuring the quality of approximation by welfare loss associated with inaccuracy of an approximate solution. Backward accuracy measures have the same shortcoming as the analysis of residuals, namely, they assess accuracy in terms of the model's parameters and do not always provide a simple way to make inferences about the size of approximation errors in the model's variables.

Our lower bound error analysis is in line with recent trends in science. In particular, National Research Council (2012) of the US National Academy of Sciences published a report with a research agenda on assessing the reliability of computational methods. The report argues that understanding of computational results can be obtained from three interrelated processes, verification, validation and uncertainty quantification (VVUQ).[1] Two goals of VVUQ that are emphasized by National Research Council (2012, p S-5) include: "... development of goal-oriented a posteriori error-estimation methods ..." and "... development of methods to estimate error bounds ...". This is precisely what we do in the paper.

The rest of the paper is organized as follows: In Section 2, we introduce the framework of lower-error bound analysis and illustrate it with examples. In Sections 3 and 4, we perform the lower error-bound analysis for a neoclassical growth model and for a new Keynesian model with Calvo pricing, respectively. In Section 5, we conclude.

[1] *Verification* refers to the process of determining how accurately a computer code solves the equations of the mathematical model. *Validation* is the process of determining the degree to which the model is an accurate representation of the real world. *Uncertainty quantification* is the process of determining how precise the results are, given uncertainties in the model and computations.

2 A lower bound on approximation errors

We introduce a framework for constructing lower bounds on approximation errors, and we discuss the relation of the lower bound error analysis to accuracy measures used in the related literature.

2.1 Testing the hypothesis that a numerical solution is accurate

We consider a system of n (possibly, nonlinear) equations with n unknowns:

$$G_i(x_1, ..., x_n) = 0, \qquad i = 1, ..., n, \tag{1}$$

or in vector notations, we have $G(x) = 0$, where $G : \mathbb{R}^n \to \mathbb{R}^n$, $n \geq 1$. (This system represents a collection of model's equations and may include a Bellman equation, Euler equations, market clearing conditions, budget constraints and laws of motion for exogenous and endogenous shocks).

Let $x^* \in \mathbb{R}^n$ and $\widehat{x} \in \mathbb{R}^n$ be the exact and approximate solutions to system (1), respectively (we assume that $\widehat{x} \neq 0$). We define an approximation error as a compensation $\delta^* \in \mathbb{R}^n$ that is needed to make an approximate solution $\widehat{x}$ to satisfy the model's equations exactly,

$$G(\widehat{x}(\mathbf{1} + \delta^*)) = 0, \tag{2}$$

where $\mathbf{1} \in \mathbb{R}^n$ is a vector of ones. Systems of equations studied in economics are often complex and finding an exact value of δ^* satisfying (2) is infeasible. (In fact, if we were able to find such a value, we would also be able to find an exact solution x^* using $x^* = \widehat{x}(\mathbf{1} + \delta^*)$).

In the paper, we propose a technique for constructing a lower bound on δ^* for those complex cases. As a first step, we remove $n - m$ equations from system (1), $1 \leq m < n$ that are most complex to solve. Without a loss of generality, we assume that we removed the last $n - m$ equations, and we denote the reduced system by $g \equiv [G_1, ..., G_m]$:

$$g_i(x_1, ..., x_n) = 0, \qquad i = 1, ..., m. \tag{3}$$

Consider now a problem of finding an approximation error δ that satisfies the reduced system of equations $g(\widehat{x}(\mathbf{1} + \delta)) = 0$. By construction, the reduced system (3) is underdetermined (it has n equations and m unknowns, $m < n$), and thus, it has multiple solutions (effectively a solution to (3) is a manifold). Consequently, there are multiple compensations δ that make an approximation $\widehat{x}$ to satisfy (3)

exactly. Let us denote a set of all possible compensations satisfying (3) for a given approximate solution $\widehat{x}$ by

$$\Omega \equiv \{\delta \in \mathbb{R}^n : g(\widehat{x}(\mathbf{1}+\delta)) = 0\}. \tag{4}$$

Our next step is to choose the smallest possible compensation $\widehat{\delta} \in \Omega$ with respect to a given norm $\|\cdot\|$, i.e.,

$$\min_{\delta \in \mathbb{R}^n} \quad \|\delta\| \qquad \text{s.t.} \qquad g(\widehat{x}(\mathbf{1}+\delta)) = 0. \tag{5}$$

The following proposition shows that the smallest possible compensation $\widehat{\delta}$ in the constructed minimization problem can never be larger than compensation δ^* in the original problem.

Proposition 1 *For a given $\widehat{x}$ and a given norm $\|\cdot\|$, we have $\| \widehat{\delta} \| \leq \|\delta^*\|$, where δ^* and $\widehat{\delta}$ are defined by (2) and (5), respectively.*

Proof. Since the reduced system (3) is a sub-system of the original system (1), any compensation δ^* satisfying (2) must be in the set of solutions Ω. Hence, there are two possibilities: i) δ^* is a solution to (5) in which case $\widehat{\delta} = \delta^*$ and hence, we have $\| \widehat{\delta} \| \leq \|\delta^*\|$ or ii) δ^* is not a solution to (5) in which case $\widehat{\delta}$ cannot not be larger than δ^* since it is the smallest possible element Ω with respect to a given norm. Both possibilities imply the claim of Proposition 1. ■

Proposition 1 allows us to form and test a hypothesis that a numerical solution is accurate. By construction $\widehat{\delta}$ is a lower bound on the actual approximation error δ^*. If even a lower bound $\widehat{\delta}$ is unacceptably large, we can reject the hypothesis that a numerical solution $\widehat{x}$ is accurate since the actual errors δ^* can never be smaller than their lower bound $\widehat{\delta}$.

The constructed lower bound depends on specific $n - m$ equations that we removed from system (1). By removing different sets of equations from the original system, we obtain different lower bounds. To show that a numerical solution is inaccurate, it is sufficient to show that any of such lower bounds is unacceptably large.

An important question is which equations should be removed from system (1). There are two considerations here: from one side, we want to remove as few equations as possible (to make the lower error bound close to the true approximation error) and from the other side, we need to remove sufficiently many equations to guarantee that a solution to optimization problem (5) can be constructed very accurately (otherwise, approximation errors in error bounds may invalidate our inferences about accuracy). In Section 3, we illustrate the construction of the reduced system by way of examples.

A convenient choice for problem (5) is an L_2 norm since it allows us to use first-order conditions (FOC), namely, we find the smallest compensation $\widehat{\delta}$ by solving the following least-squares problem:

$$\min_{\delta \in \mathbb{R}^n} \quad \delta^\top \delta \quad \text{s.t.} \quad g\left(\widehat{x}\left(\mathbf{1}+\delta\right)\right) = 0. \tag{6}$$

A necessary condition for the existence of a local minimum $\widehat{\delta}$ in (6) follows by a version of the well-known Theorem of Lagrange: (i) $g\left(\widehat{x}\left(\mathbf{1}+\widehat{\delta}\right)\right)$ must be full ranked in a neighborhood of $\widehat{x}\left(\mathbf{1}+\widehat{\delta}\right)$; and (ii) $\widehat{x}\left(\mathbf{1}+\widehat{\delta}\right)$ must be a critical point of the Lagrange function, $\delta^\top \delta + \lambda g\left(\widehat{x}\left(\mathbf{1}+\delta\right)\right)$, where $\lambda \in \mathbb{R}^m$ is a vector of Lagrange multipliers, i.e.,

$$2\widehat{\delta} + \lambda \nabla g\left(\widehat{x}\left(\mathbf{1}+\widehat{\delta}\right)\right)\widehat{x} = 0, \tag{7}$$

where ∇g denotes a gradient of g. Furthermore, a sufficient condition for a local minimum is that the Lagrangian function is convex on a subset of $\mathbb{R}^n$ defined by $Z\left(\widehat{\delta}\right) = \left\{z \in \mathbb{R}^n : \nabla g\left(\widehat{x}\left(\mathbf{1}+\widehat{\delta}\right)\right) z = 0\right\}$; see, e.g., Sundaram (1996, Theorems 5.1 and 5.4) for proofs of these results.

Instead of L_2, we can use other norms for measuring compensations, for example, a least absolute deviation L_1 or a maximum error L_∞. Furthermore, in some economic applications, we can tolerate large approximation errors in some variables but we need very accurate solutions in other variables. In this case, the approximation errors can be weighted by a measure of their economic significance in the objective function. For example, the objective function in (6) can be modified to $\delta^\top W \delta$, where W is an $n \times n$ matrix of weights (this case is similar to a weighted least-squares in econometrics).

2.2 Two-dimensional case

We now illustrate a construction of a lower bound on approximation errors in a two-dimensional case. Let (x_1^*, x_2^*) and $(\widehat{x}_1, \widehat{x}_2)$ denote the exact and approximate solutions to a two-dimensional version of the system (1), namely, $G_i(x_1, x_2) = 0$, $i = 1, 2$ (again, we assume that $(\widehat{x}_1, \widehat{x}_2) \neq 0$). Following (2), we define the approximation error $\left(\delta_{x_1}^*, \delta_{x_2}^*\right)$ by $G_i\left(x_1^*\left(1+\delta_{x_1}^*\right), x_2^*\left(1+\delta_{x_1}^*\right)\right) = 0$, $i = 1, 2$.

To construct a lower bound on approximation errors, we remove equation $G_2(x_1, x_2) = 0$ and we focus on the reduced system composed of just one equation $g(x_1, x_2) \equiv G_1(x_1, x_2) = 0$. Following (4), we define a set of compensations Ω that are consistent with a restriction g:

$$\Omega \equiv \left\{ (\delta_{x_1}, \delta_{x_2}) \in \mathbb{R}^2 : g\left(\underbrace{\widehat{x}_1\left(1+\delta_{x_1}\right)}_{=x_1^*}, \underbrace{\widehat{x}_2\left(1+\delta_{x_2}\right)}_{=x_2^*} \right) = 0 \right\}. \tag{8}$$

As we mentioned earlier, the reduced system of equations g is underdetermined and there are multiple compensations δ_{x_1} and δ_{x_2} that are consistent with (8). As an illustration, consider a special case when g is linear, i.e.,

$$g(x_1, x_2) = a_1 x_1 + a_2 x_2, \tag{9}$$

where a_1 and a_2 are constant coefficients. To describe all compensations satisfying (8), we can fix any δ_{x_1}, and we can find δ_{x_2} from (8) using (9) as follows:

$$\delta_{x_2} = \frac{a_1 \widehat{x}_1}{a_2 \widehat{x}_2} (1 + \delta_{x_1}) - 1. \tag{10}$$

From all possible compensations satisfying (10), we select the smallest one with respect to the least-squares norm by solving a two-dimensional version of the least-squares problem (6):

$$\min_{\delta_{x_1}, \delta_{x_2}} \quad \delta_{x_1}^2 + \delta_{x_2}^2 \tag{11}$$

$$\text{s.t. } g(\widehat{x}_1 (1 + \delta_{x_1}), \widehat{x}_2 (1 + \delta_{x_2})) = 0. \tag{12}$$

An interior solution of (11), (12) satisfies

$$\frac{\delta_{x_1}}{\delta_{x_2}} = \frac{g_{x_1}(\widehat{x}_1 (1 + \delta_{x_1}), \widehat{x}_2 (1 + \delta_{x_2}))\, \widehat{x}_1}{g_{x_2}(\widehat{x}_1 (1 + \delta_{x_1}), \widehat{x}_2 (1 + \delta_{x_2}))\, \widehat{x}_2}. \tag{13}$$

Hence, to construct the smallest possible approximation errors, we must solve a system of two equations (12), (13) with respect to two unknowns δ_{x_1} and δ_{x_2}.

For the case of a linear equation (9), we can solve this system in a closed form,

$$\widehat{\delta}_{x_i} = - \frac{a_i \widehat{x}_i (a_1 \widehat{x}_1 + a_2 \widehat{x}_2)}{(a_1 \widehat{x}_1)^2 + (a_2 \widehat{x}_2)^2}, \qquad i = 1, 2, \tag{14}$$

where to derive (14), we used the fact that $g_{x_1}(\cdot) = a_1$ and $g_{x_2}(\cdot) = a_2$.

However, for a general nonlinear restriction $g(x_1, x_2) = 0$, system (12), (13) does not admit a closed form representation. If approximation errors are small, a sufficiently accurate solution to (12), (13) can be obtained by using a first-order Taylor expansion:

$$g(\widehat{x}_1 (1 + \delta_{x_1}), \widehat{x}_2 (1 + \delta_{x_2})) \approx g(\widehat{x}_1, \widehat{x}_2) + g_{x_1}(\widehat{x}_1, \widehat{x}_2)\, \widehat{x}_1 \delta_{x_1} + g_{x_2}(\widehat{x}_1, \widehat{x}_2)\, \widehat{x}_2 \delta_{x_2}. \tag{15}$$

Combining (15) with FOC (13) evaluated in $(\widehat{x}_1, \widehat{x}_2)$ yields:

$$\widehat{\delta}_{x_i} = - \frac{g_{x_i}(\widehat{x}_1, \widehat{x}_2)\, \widehat{x}_i g(\widehat{x}_1, \widehat{x}_2)}{[g_{x_1}(\widehat{x}_1, \widehat{x}_2)]^2 (\widehat{x}_1)^2 + [g_{x2}(\widehat{x}_1, \widehat{x}_2)]^2 (\widehat{x}_2)^2}, \qquad i = 1, 2. \tag{16}$$

If approximation (16) is not sufficiently accurate, we need to either construct a Taylor expansions of a higher order or to find a non-linear solution to (12), (13) using a numerical solver such as a Newton method. In that case, a linear approximation (16) can be used as an initial guess for a numerical solver.

2.3 Relation of lower bound error analysis to other notions of approximation errors in the literature

There are three main approaches to accuracy evaluation in economic literature: a forward error analysis, an analysis of residuals and a backward error analysis. A forward error analysis assesses an approximation error in the solution of a given model. (Hence, our lower bound error analysis is a variant of a forward error analysis). Analysis of residuals consists in evaluating residuals in model's equations for a given approximate solution. Finally, a backward error analysis proceed in a reverse manner: it takes an approximate solution as given and asks how much the model itself must be modified to make an approximate solution to satisfy all the model's equations. Below we discuss the relation of these three approaches to our lower bound error analysis.

2.3.1 A conventional forward error analysis

A conventional forward error analysis aims to construct an upper bound on the size of the approximation errors, see, e.g., Bertsekas and Tsitsiklis (1996) , Santos and Vigo-Aguillar (1998), Santos (2000), Schmitt-Grohé and Uribe (2004), and Santos and Peralta-Alva (2005), among others; see Santos and Peralta-Alva (2014) for a review of this literature. The upper error bound corresponds to a pessimistic – worst case – scenario. The following question is addressed: What are the largest possible approximation errors that are consistent with a given numerical solution? The upper bound error analysis provides a sufficient condition for accuracy: If an upper bound on approximation errors is still acceptable, we conclude that an approximate solution is sufficiently accurate since the actual errors can never be larger.

In contrast, our lower-error bound analysis focuses on optimistic – best-case – scenarios. Here, we ask: How small approximation errors can potentially be made if we allow to violate some of the model's equations? If the resulting lower error bound is still unacceptably large, we conclude that a numerical solution is inaccurate since the actual approximation errors can never be smaller. Hence, our lower bound error analysis provides a necessary condition for accuracy.

The upper and lower error-bound analyses are analogous to necessary and sufficiency tests in econometrics. A lower bound error test can be used to reject the hypothesis that a numerical solution is accurate (when the errors are large even under the most optimistic view) but not to accept it (because errors that are small under an optimistic view can be unacceptably large under a pessimistic view). To accept the hypothesis that a numerical solution is accurate, we need more stringent accuracy tests such as an upper bound error analysis. In practice, upper error bound tests tend to be too pessimistic and they reject many accurate

solutions, while our lower bound error test can be too optimistic and may fail to reject inaccurate solutions. We view the upper and lower bound error analysis as complementary.

Finally, an important limitation of conventional forward error analysis is that it is restricted to dynamic programing problems. The construction of upper bounds in the literature relies on the fact that the Bellman operator is a contraction mapping. This kind of error analysis is not directly applicable to non-optimal equilibrium problems such as a new Keynesian model studied in the present paper. In turn, our lower bound error analysis is applicable to both optimal dynamic programing problems and non-optimal equilibrium problems.

2.3.2 An analysis of residuals in the model's equations

A commonly used accuracy measure in the literature is residuals in the model's equations (such as a Bellman equation, Euler equations, market clearing conditions, budget constraints and laws of motion for exogenous and endogenous shocks); see, e.g., Judd (1992), Jin and Judd (2000), Aruoba et al. (2005), Juillard and Villemot (2011), Judd et al. (2011a); and also, see a statistical test of residuals by Den Haan and Marcet (1994).

We define residuals in a unit-free way. First, we rewrite equations $G_i(x) = 0$, $i = 1, ..., n$, of the system (1) in the form $x_i = H_i(x_{-i})$, $i = 1, ..., n$, where x_{-i} denotes a set of all model's variables with an exception of x_i (we assume that it is possible to do). Then, we define residuals $\mathcal{R} \equiv (\mathcal{R}^1, ..., \mathcal{R}^n)$ in the model's equations of system (1) to satisfy

$$\widehat{x}_i \left(1 + \mathcal{R}^i\right) \equiv H_i(\widehat{x}_{-i}), \qquad i = 1, ..., n. \tag{17}$$

Under definition (17), a residual $\mathcal{R}^i$ shows how large would an approximation error in a given variable x_i be if we assume that approximation errors in all other variables x_{-i} are zeros. Even though $\mathcal{R}$ resembles the true approximation errors δ^* defined in (2), there is an important difference: δ^* is constructed in a way that is consistent with all model's equations, while $\mathcal{R}$ is not. For example, to compute $\mathcal{R}^1$ in the two-dimensional case, we assume that x_2 is computed without errors, while to compute $\mathcal{R}^2$, we assume that x_1 is computed without errors, which cannot be true simultaneously.

The following proposition shows that residuals in the model's equations constitute neither upper nor lower bounds on approximation errors.

Proposition 2 *For* $\delta^* = \left(\delta^*_{x_1}, ..., \delta^*_{x_n}\right)$ *and* $\mathcal{R} = (\mathcal{R}^1, ..., \mathcal{R}^n)$ *defined in (2) and (17), respectively, we have* $\left|\delta^*_{x_1}\right| \gtreqless |\mathcal{R}^i|$ *for* $i = 1, ..., n$.

Proof. We prove the proposition by way of example. Consider a system of linear equations

$$a_1 x_1 + a_2 x_2 = 1, \tag{18}$$
$$b_1 x_1 + b_2 x_2 = 1, \tag{19}$$

where a_1, a_2, b_1, b_2 are constant coefficients. According to definition (2), the approximation errors $\delta^*_{x_1}$ and $\delta^*_{x_2}$ satisfy

$$a_1 \widehat{x}_1 \left(1 + \delta^*_{x_1}\right) + a_2 \widehat{x}_2 \left(1 + \delta^*_{x_2}\right) = 1, \tag{20}$$
$$b_1 \widehat{x}_1 \left(1 + \delta^*_{x_1}\right) + b_2 \widehat{x}_2 \left(1 + \delta^*_{x_2}\right) = 1. \tag{21}$$

Following (17), we construct residuals $\mathcal{R}^1$ and $\mathcal{R}^2$ as

$$a_1 \widehat{x}_1 \left(1 + \mathcal{R}^1\right) + a_2 \widehat{x}_2 = 1, \tag{22}$$
$$b_1 \widehat{x}_1 + b_2 \widehat{x}_2 \left(1 + \mathcal{R}^2\right) = 1. \tag{23}$$

By combining (20)–(23), we get

$$\delta^*_{x_1} + v_1 \delta^*_{x_2} = \mathcal{R}^1, \tag{24}$$
$$v_2 \delta^*_{x_1} + \delta^*_{x_2} = \mathcal{R}^2, \tag{25}$$

where $v_1 \equiv \frac{a_2 \widehat{x}_2}{a_1 \widehat{x}_1}$ and $v_2 \equiv \frac{b_1 \widehat{x}_1}{b_2 \widehat{x}_2}$. Without a loss of generality, let us assume $v_1 > 0$ and $v_2 > 0$, and let us analyze a relation between $\mathcal{R}^1$ and $\delta^*_{x_1}$. If $\delta^*_{x_1}$, $\delta^*_{x_2} > 0$, then we have $|\mathcal{R}^1| > \left|\delta^*_{x_1}\right|$. Furthermore, if $\delta^*_{x_1} > 0$, $\delta^*_{x_2} < 0$ and $\delta^*_{x_1} + v_1 \delta^*_{x_2} > -\delta^*_{x_1}$, then we have $|\mathcal{R}^1| < \left|\delta^*_{x_1}\right|$. Finally, if $\delta^*_{x_2} = 0$, then we have $|\mathcal{R}^1| = \left|\delta^*_{x_1}\right|$. ■

Proposition 2 shows that residuals and approximation errors cannot be ranked, namely, residuals can be either larger or smaller or equal to approximation errors in absolute value (even if we define them in a comparable way). Moreover, the difference between residuals and approximation errors can be made arbitrary large by varying the coefficients a_1, a_2, b_1, b_2 in the system (18), (19). Thus, small residuals in the model's equations do not necessarily imply small approximation errors. The relation between the residuals and approximation errors is established in the literature only for a special case of strongly concave infinite-horizon optimization problems by Santos (2000) who shows that approximation errors in policy functions are of the same order of magnitude as the size of the Euler equation residuals. In general, such a relation is not known.

An example used in the proof of Proposition 2 has another interesting implication, namely, if $\delta^*_{x_1} = -v_1 \delta^*_{x_2}$, then the residual in (24) is zero $|\mathcal{R}^1| = 0$. Thus, zero residuals in some model's equation are consistent with arbitrary large approximation errors as long as these errors have offset one another.

Finally, the residual measure of accuracy (17) has another undesirable feature, namely, it is not invariant to reordering of equations and variables in (1). For example, we would obtain different residuals if instead of expressing x_1 and x_2 from $G_1(x) = 0$ and $G_2(x) = 0$, respectively, we express x_1 and x_2 from $G_2(x) = 0$ and $G_1(x) = 0$, respectively. In contrast, our lower error bounds are independent of a specific way in which the system (1) is written.

2.3.3 A backward and mixed forward-backward error analysis

A backward error analysis introduced in Wilkinson (1963) poses the following question: How much the parameters of a model must be modified in order to make an approximate solution to satisfy the model's equation exactly? A mixed forward-backward analysis introduced in Higham (1996) is an extension of backward analysis which allows for changes in both equilibrium quantities and the model's parameters. Sims (1990) proposed an accuracy test which is similar in spirit to the backward error analysis: he measures accuracy by how far the distribution of the true stochastic shocks is situated from the distribution of stochastic shocks that is implied by the approximate solution. Kubler and Schmedders (2005) shows how a backward and a mixed backward-forward analyses can be used to evaluate the accuracy of numerical solutions in a life-cycle model with incomplete markets and heterogeneous agents. Finally, Kogan and Mitra (2013) propose to measure the quality of approximation in terms of a welfare loss that results from inaccuracy of an approximate solution. They construct a supplementary model with perfect foresight and assess the difference in welfare between that supplementary model and the true stochastic model with an approximate solution–this provides an upper bound on the welfare loss.

A backward error analysis can be incorporated in the framework of Section 2.1 as follows: Let us assume that G in (1) depends not only on x but also on a vector of parameters b, i.e., the original system modifies to $G(x; b) = 0$, where $G : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$. We then ask: How much must the parameter vector b be modified to make a given approximate solution $\widehat{x}$ to become an exact solution, i.e., we have $G\left(\widehat{x}; \widehat{b}\right) = 0$. Provided that we find a vector of parameters $\widehat{b}$ that satisfies the latter condition, we measure the accuracy by the distance $\left\|\widehat{b} - b\right\|$.

The backward and mixed forward-backward accuracy measures are also indirect measures of accuracy and are generally subject to the same critique as the analysis of residuals. Namely, they do not show the distance between the true and approximate solutions, x and $\widehat{x}$, but the distance between the parameters b and $\widehat{b}$ or some mixture of the parameters and solutions. Furthermore, it is not clear whether or not for any given model and for any approximate solution, one can find a supplementary model (parameter vector $\widehat{b}$) that leads to zero approximation er-

rors. In Section 3.4, we illustrate the construction of backward accuracy measures in the context of a neoclassical stochastic growth model.

3 Application 1: the optimal growth model

In this section, we show how to construct lower bounds on approximation errors for the standard neoclassical stochastic growth model. We then assess such bounds for numerical solutions produced by a first- and second-order perturbation methods.

3.1 The model

The representative agent solves

$$\max_{\{k_{t+1},c_t\}_{t=0,\ldots,\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\right) \tag{26}$$

$$\text{s.t. } c_t + k_{t+1} = (1-d)\, k_t + \exp\left(\theta_t\right) A f\left(k_t\right), \tag{27}$$

$$\theta_{t+1} = \rho\theta_t + \epsilon_{t+1}, \qquad \epsilon_{t+1} \sim \mathcal{N}\left(0, \sigma^2\right), \tag{28}$$

where (k_0, θ_0) is given; E_t is the expectation operator conditional on information at time t; c_t, k_t and θ_t are consumption, capital and productivity level, respectively; $\beta \in (0,1)$ is the discount factor; $d \in (0,1]$ is the depreciation rate of capital; $\rho \in (-1,1)$ is the autocorrelation coefficient of the productivity level; $\sigma \geq 0$ is the standard deviation of the productivity shock; $A > 0$ is a normalizing constant in output; u and f are strictly increasing, continuously differentiable and concave; u' and f' denote the first derivatives of u and f, respectively.

The Euler equation of (26)–(28) is

$$u'\left(c_t\right) = \beta E_t\left\{u'\left(c_{t+1}\right)\left[1 - d + \exp\left(\theta_{t+1}\right) A f'\left(k_{t+1}\right)\right]\right\}. \tag{29}$$

A solution to the model is policy functions $c_t = C\left(k_t, \theta_t\right)$ and $k_{t+1} = K\left(k_t, \theta_t\right)$ that satisfy (27), (28) and (29) for all (k_t, θ_t) within the relevant domain.

3.2 State-contingent approximation errors

We first show a definition of approximation errors in the sense (2). Let us consider a numerical solution to (26)–(28) in the form of an approximation to consumption and capital functions, $\widehat{C} \approx C$ and $\widehat{K} \approx K$, respectively.

We define approximation errors δ_C and δ_K as state contingent functions satisfying the model's equations (27) and (29):

$$\widehat{C}\left(k,\theta\right)\left(1 + \delta_C\left(k,\theta\right)\right) + \widehat{K}\left(k,\theta\right)\left(1 + \delta_K\left(k,\theta\right)\right) = (1-d)\, k + \exp\left(\theta\right) A f\left(k\right), \tag{30}$$

$$u'\left(\widehat{C}\left(k,\theta\right)\left(1+\delta_C\left(k,\theta\right)\right)\right)=\beta E_t\left\{u'\left(\widehat{C}\left(k',\theta'\right)\left(1+\delta_C\left(k',\theta'\right)\right)\right)\times \left[1-d+\exp\left(\theta'\right)Af'\left(\widehat{K}\left(k,\theta\right)\left(1+\delta_K\left(k,\theta\right)\right)\right)\right]\right\}. \quad (31)$$

Finding δ_C and δ_K exactly from (30) and (31) is generally infeasible. It is possible to compute δ_C and δ_K numerically up to some degree of accuracy. However, if the error functions δ_C and δ_K are constructed with errors themselves, we would not be able to tell whether such functions measure the accuracy of numerical solution $\widehat{C}$ and $\widehat{K}$ or they measure the errors in their own approximation $\widehat{\delta}_C \approx \delta_C^*$ and $\widehat{\delta}_K \approx \delta_K^*$. That is, having approximation errors in approximation errors would contaminate the error bound analysis and invalidate the accuracy inferences.

3.3 A lower bound on approximation error

Given that a construction of state contingent error functions is infeasible, we focus on constructing their lower bounds.

3.3.1 Defining a lower error bound

We evaluate approximation errors in equilibrium quantities in a point-by-point manner without exploiting a state-contingent structure of the error functions.

We first write budget constraint (27) for $t \geq 0$ as

$$\underbrace{\widehat{c}_t\left(1+\delta_{c_t}\right)}_{=c_t}+\underbrace{\widehat{k}_{t+1}\left(1+\delta_{k_{t+1}}\right)}_{=k_{t+1}}=\exp\left(\theta_t\right)Af\left(k_t\right)+\left(1-d\right)k_t, \quad (32)$$

where δ_{c_t} and $\delta_{k_{t+1}}$ are the approximation errors that show how much an approximate solution $\widehat{c}_t$ and $\widehat{k}_{t+1}$ must be modified to become the true solution c_t and k_{t+1}, respectively.

We then represent Euler equation (29) for $t \geq 0$ as

$$u'\underbrace{\left(\widehat{c}_t\left(1+\delta_{c_t}\right)\right)}_{=c_t}=\beta E_t\left\{u'\underbrace{\left(\widehat{c}_{t+1}\left(1+\delta_{c_{t+1}}\right)\right)}_{=c_{t+1}}\times \left[1-d+\exp\left(\rho\theta_t+\epsilon_{t+1}\right)Af'\underbrace{\left(\widehat{k}_{t+1}\left(1+\delta_{k_{t+1}}\right)\right)}_{=k_{t+1}}\right]\right\}, \quad (33)$$

where δ_{c_t} and $\delta_{k_{t+1}}$ are defined as in (32); and $\delta_{c_{t+1}}$ represents an approximation error in calculating the conditional expectation, namely, it shows how much the

approximate solution $\widehat{c}_{t+1}$ must be modified on average to make the Euler equation (33) to be satisfied exactly (under given $\delta_{k_{t+1}}$). Hence, our error analysis disregards the fact that δ_{c_t} and $\delta_{c_{t+1}}$ are given by the same state contingent function defined by (30) and (31), i.e., $\delta_{c_t} = \delta_C(k_t, \theta_t)$ and $\delta_{c_{t+1}} = \delta_C(k_{t+1}, \theta_{t+1})$, respectively. Instead, we will choose δ_{c_t} and $\delta_{c_{t+1}}$ in the way that makes them as small as possible subject to the constraints (32), (33).

System (32), (33) is underdetermined and does not identify δ_{c_t}, $\delta_{k_{t+1}}$ and $\delta_{c_{t+1}}$ uniquely (we have two equations with three unknowns). In particular, fixing arbitrary one out three unknowns, let us say $\delta_{k_{t+1}}$, enables us to find the remaining unknowns δ_{c_t} and $\delta_{c_{t+1}}$ to satisfy (32), (33) exactly (in other words, a solution to (32), (33) is a manifold). To construct a lower error bound, we solve the least-squares problem of type (6):

$$\min_{\delta_{c_t}, \delta_{k_{t+1}}, \delta_{c_{t+1}}} \quad \delta_{c_t}^2 + \delta_{k_{t+1}}^2 + \delta_{c_{t+1}}^2 \quad \text{s.t.} \quad (32), (33). \tag{34}$$

Problem (34) determines the lower error bound and provides a necessary condition for accuracy. If the lower error bound produced by (34) is still large, the actual errors given by the state contingent functions $\delta_{c_t} = \delta_C(k_t, \theta_t)$ and $\delta_{c_{t+1}} = \delta_C(k_{t+1}, \theta_{t+1})$ must be even larger, and we conclude that a numerical solution is inaccurate.

3.3.2 Numerical experiments

We use Dynare to compute the first- and second-order perturbation solutions, referred to as *PER1* and *PER2*, respectively; for a description of this software, see Adjemian et al. (2011). We parameterize the model (26)–(28) by assuming $u(c_t) = \frac{c_t^{1-\gamma}-1}{1-\gamma}$ with $\gamma \in \left\{\frac{1}{10}, 1, 10\right\}$ and $f(k_t) = k_t^{\alpha}$ with $\alpha = 0.33$. We set $\beta = 0.99$, $d = 0.025$, $\rho = 0.95$ and $\sigma = 0.01$, and we normalize the steady state of capital to one by assuming $A = \frac{1/\beta-(1-d)}{\alpha}$.

To simulate a time series solution, we draw a sequence of shocks $\{\epsilon_\tau\}_{\tau=1}^{\mathcal{T}}$ from $\mathcal{N}(0, \sigma^2)$ with $\mathcal{T} = 10,200$, construct $\{\theta_\tau\}_{\tau=0}^{\mathcal{T}}$ using (28), and compute the series of capital and consumption using the perturbation solutions (we disregard the first 200 observations to eliminate the effect of initial conditions). To compute expectation in (33), we use a 10-point Gauss-Hermite quadrature integration rule.

We solve minimization problem (34) numerically for each state (k_t, θ_t) realized on a stochastic simulation. To find initial guesses for δ_{c_t}, $\delta_{k_{t+1}}$, $\delta_{c_{t+1}}$, we compute first-order Taylor expansions of (32), (33) around $\delta_{c_t} \to 0$, $\delta_{k_{t+1}} \to 0$, $\delta_{c_{t+1}} \to 0$, and we minimize the squared sum of errors subject to the obtained linear constraints (the resulting problem is a quadratic programming problem); see Appendix A1 for details. We subsequently use a Newton solver to compute a highly accurate

nonlinear solution (34) using the first-order approximation as an initial guess; see Appendix A2 for details. The results are provided in Table 1.

Table 1: Approximation errors in the equilibrium allocations in the neoclassical stochastic growth model

	$\gamma = \frac{1}{10}$			$\gamma = 1$			$\gamma = 10$		
Errors	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$
PER1									
L_1	-3.96	-4.07	-3.94	-4.80	-4.11	-4.65	-4.35	-3.75	-4.44
L_∞	-2.90	-2.98	-2.88	-4.02	-3.04	-3.77	-3.55	-2.62	-3.88
PER2									
L_1	-5.63	-5.75	-5.62	-6.30	-5.68	-6.25	-5.57	-4.75	-5.59
L_∞	-4.53	-4.39	-4.52	-5.15	-4.43	-4.99	-4.42	-3.65	-4.28

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of the lower bounds on approximation errors across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observations; and γ is the coefficient of risk aversion.

Across all the cases, highest maximum approximation errors are $10^{-2.62} \approx 0.25\%$ and $10^{-3.65} \approx 0.025\%$ for *PER1* and *PER2*, respectively, which corresponds to the case of a large risk aversion coefficient, $\gamma = 10$. These numbers are sufficiently low, which allows us to conclude that if we take an optimistic view, the approximation errors are acceptable in size. Again, our test is a necessary condition for accuracy and does not allow us to conclude that perturbation solutions are accurate. We can only say that we cannot reject the hypothesis that perturbation methods are sufficiently reliable for the standard growth model on the basis of our numerical experiments.

3.3.3 Modifications and extensions of lower bound error analysis

We now discuss possible modifications and extensions of the constructed accuracy measures.

Choice of error functions to construct. There are many possible ways to define approximation errors. First, we could consider errors in other model's variables, for example, errors in the investment or output functions instead of those in capital or consumption functions. Second, there are different ways of modeling approximation errors in conditional expectations, for example, we can represent errors in Euler equation (29) as

$$u'\left(\widehat{c}_t\left(1+\delta_{c_t}\right)\right) = \underbrace{\beta\left(1+\delta_{E_t}\right)\widehat{E}_t}_{=E_t[u'(c_{t+1})(1-d+\exp(\rho\theta_t+\epsilon_{t+1})Af'(k_{t+1}))]}, \tag{35}$$

where δ_{E_t} is an approximation error in conditional expectation function $E_t[\cdot]$. We can use a new condition (35) as a restriction in the least-squares problem (34) instead of (33) by changing the objective function to $\delta_{c_t}^2 + \delta_{k_{t+1}}^2 + \delta_{E_t}^2$.

A potential shortcoming of this alternative representation is that the error in $E_t[\cdot]$ depends on a specific representation of the utility function, so that $\delta_{c_t}^2$, $\delta_{k_{t+1}}^2$, $\delta_{E_t}^2$ are not expressed in comparable units, and introducing a trade-off between these errors in the objective function may not lead to a meaningful accuracy criteria. In contrast, our baseline representation $\delta_{c_t}^2 + \delta_{k_{t+1}}^2 + \delta_{c_{t+1}}^2$ in (33) contains approximation errors in comparable quantities and is not subject to this shortcoming. Kubler and Schmedders (2005) also measure the error in the conditional expectation function δ_{E_t} by the average adjustment to the future consumption $\delta_{c_{t+1}}$ as is done under our baseline representation (33).

Weighting approximation errors by their relative importance. In applications, one may want to approximate some variables more accurately than others. For example, in a model, aimed at explaining consumption growth, one may want to get a more accurate approximation for consumption function than for capital function. This case can be incorporated into our analysis by modifying (34) to minimize a weighted sum of squared approximation errors:

$$\min_{\delta_{c_t}, \delta_{k_{t+1}}, \delta_{c_{t+1}}} \quad w_1\delta_{c_t}^2 + w_2\delta_{k_{t+1}}^2 + w_3\delta_{c_{t+1}}^2 \quad \text{s.t.} \quad (32), (33), \tag{36}$$

where $w_1, w_2, w_3 \in [0, \infty)$ are given weights. By changing the weights, one can vary the relative importance attached to errors in different variables. We can then ask the following question: How much error can be hided at most in variables that we care less about? We perform experiments in which we set one of the three approximation errors to zero (equivalently, we set its weight to a very large number), and we then compute the remaining two errors; see Appendix A3 for a detailed description of these experiments.

We find that if a lower error bound on either current or future consumption is restricted to be zero, the resulting lower bounds are very similar to the unrestricted bounds in Table 1. However, if we set $\delta_{k_{t+1}} = 0$, our best-case scenario worsens considerably and a lower error bound increases by an order of magnitude to $10^{-1.65} \approx 2\%$ and $10^{-2.65} \approx 0.2\%$ for *PER1* and *PER2*, respectively; see Tables 7, 8 and 9 in Appendix A3 for these results. That is, if we want to believe that our capital function is computed very accurately, we must accept the fact that our consumption function is quite inaccurate.

3.4 Alternative accuracy measures for the growth model

In Section 2.3, we described three alternative accuracy measures in the literature: a forward error analysis, an analysis of residuals and a mixed forward-backward error analysis. We now illustrate these accuracy measures in the context of the studied growth model.

A forward error analysis. An upper bound forward error analysis for the standard growth model is carefully implemented in Santos (2000). His analysis relies on the fact that the standard growth model (26)–(28) can be reformulated as a dynamic programing problem. A contraction mapping property of the Bellman operator makes it possible to construct the upper bound on errors analytically.

The analysis of Santos (2000) shows that a worst-case scenario can often be too pessimistic and may lead to a rejection of numerical solutions that are sufficiently accurate. For example, under the standard calibration of the optimal growth model, Santos (2000, Table I) obtains an upper error bound on a policy function of order 10^3. Consequently, he shows that this error bound can be reduced by about three orders of magnitude by using some additional information from a specific numerical solution.

Analysis of residuals in model's equations. We define unit-free residuals in a point (k_t, θ_t) by re-writing (29) and (27) as follows:

$$\mathcal{R}^1(k_t,\theta_t) \equiv \frac{u'^{-1}\left[\beta E_t\left\{u'\left(\widehat{c}_{t+1}\right)\left[1-d+\exp\left(\rho\theta_t+\epsilon_{t+1}\right)Af'\left(\widehat{k}_{t+1}\right)\right]\right\}\right]}{\widehat{c}_t}-1, \tag{37}$$

$$\mathcal{R}^2(k_t,\theta_t) \equiv \frac{\exp(\theta_t)Af(k_t)+(1-d)k_t-\widehat{c}_t}{\widehat{k}_{t+1}}-1, \tag{38}$$

where $\widehat{c}_{t+1}=\widehat{C}(k_{t+1},\theta_{t+1})=\widehat{C}\left(\widehat{K}(k_t,\theta_t),\rho\theta_t+\epsilon_{t+1}\right)$, $\widehat{c}_t=\widehat{C}(k_t,\theta_t)$, and $\widehat{k}_{t+1}=\widehat{K}(k_t,\theta_t)$. Here, we express $\mathcal{R}^1$ and $\mathcal{R}^2$ in terms of consumption and capital units, respectively, which is parallel to the definitions of approximation errors δ_{c_t} and $\delta_{k_{t+1}}$ in our lower error-bound analysis. Namely, $\mathcal{R}^1(k_t,\theta_t)$ is the same as δ_{c_t} if we assume that c_{t+1} and k_{t+1} are computed without errors (i.e., we set $\delta_{k_{t+1}}=\delta_{c_{t+1}}=0$) and $\mathcal{R}^2(k_t,\theta_t)$ is the same as $\delta_{k_{t+1}}$ if we assume that c_t is computed without errors (i.e., we set $\delta_{c_t}=0$).

We compute $\mathcal{R}^1(k_t,\theta_t)$ and $\mathcal{R}^2(k_t,\theta_t)$ on a set of simulated points (for the details of the simulation procedure, see Section 3.3.2). The results are provided in Table 2.

Table 2: Residuals of the equilibrium conditions in the neoclassical stochastic growth model

Residuals	$\gamma = \frac{1}{10}$		$\gamma = 1$		$\gamma = 10$	
	R^1	R^2	R^1	R^2	R^1	R^2
PER1						
L_1	−3.61	−4.12	−4.40	−4.12	−4.09	−3.74
L_∞	−2.55	−3.02	−3.55	−3.04	−3.52	−2.61
PER2						
L_1	−5.29	−5.80	−5.96	−5.69	−5.30	−4.75
L_∞	−4.20	−4.41	−4.74	−4.44	−4.05	−3.65

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of residuals in the model's equations across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observations; and γ is the coefficient of risk aversion.

As we can see, the maximum residuals across the two equilibrium conditions are below $\frac{1}{3}\% \approx 10^{-2.55}$ for *PER1* and about $\frac{1}{45}\% \approx 10^{-3.65}$ for *PER2*. Thus, both *PER1* and *PER2* produce relatively small residuals in this model. A comparison with Table 1 shows that the maximum approximation errors are somewhat smaller than the maximum residuals. This tendency was robust in our experiments, however, it cannot be viewed as a generic property of the model and / or solution methods; in general, residuals and approximation errors cannot be ranked as an example in Section 2.3.2 shows.

A backward error analysis. There are many possible ways to implement a backward error analysis for the optimal growth model (26)–(28). We choose one such a way by measuring the accuracy in the Euler equation (29) and budget constraint (27) by the implied values of the parameters β and δ, denoted by $\beta(k_t, \theta_t)$ and $d(k_t, \theta_t)$, respectively

$$\beta(k_t, \theta_t) = E_t \left\{ \frac{u'(\widehat{c}_{t+1})}{u'(\widehat{c}_t)} \left[1 - d + \exp(\rho\theta_t + \epsilon_{t+1}) A f'\left(\widehat{k}_{t+1}\right) \right] \right\}^{-1}, \tag{39}$$

$$d(k_t, \theta_t) \equiv \left\{ 1 - \frac{\widehat{c}_t + \widehat{k}_{t+1} - \exp(\theta_t) A f(k_t)}{k_t} \right\}, \tag{40}$$

We compute $\beta(k_t, \theta_t)$ and $d(k_t, \theta_t)$ on the same set of simulated points as all our previous statistics; see Section 3.3.2. The results are provided in Table 3. The accuracy implications here are similar to those in Tables 1 and 2. The least accurate solution is obtained under $\gamma = 10$, in particular, *PER1* implies that $\beta(k_t, \theta_t)$ and $d(k_t, \theta_t)$ range within $[.9870, .9894]$ and $[.0225, .0261]$ which correspond to up to 0.3% and 10% deviations from their true values $\beta = .99$ and $d = .025$, respectively.

Table 3: The implied parameter values in the neoclassical stochastic growth model

	$\gamma = \frac{1}{10}$		$\gamma = 1$		$\gamma = 10$	
Parameters	$\beta(k_t, \theta_t)$	$d(k_t, \theta_t)$	$\beta(k_t, \theta_t)$	$d(k_t, \theta_t)$	$\beta(k_t, \theta_t)$	$d(k_t, \theta_t)$
PER1						
mean	.9900	.0251	.9900	.0251	.9892	.0249
min	.9897	.0249	.9857	.0248	.9870	.0225
max	.9901	.0259	.9900	.0259	.9894	.0261
PER2						
mean	.9900	.0250	.9900	.0250	.9900	.0250
min	.9900	.0250	.9900	.0250	.9892	.0248
max	.9900	.0250	.9900	.0250	.9909	.0251

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; "mean", "min" and "max" are, repectively, the average, minimum and maximum of the value of the corresponding model's parameter on a stochastic simulation of 10,000 observations; and γ is the coefficient of risk aversion.

PER2 is more accurate than *PER1*, in particular, under $\gamma \in \left\{\frac{1}{10}, 1\right\}$, the parameter values implies by *PER2* coincide with their true values at least up to four digits.

Our results illustrate a shortcoming of the backward error analysis, namely, it is not clear how to interpret the implied deviations in the parameters. Percentage differences in the parameter values may be not an informative statistic for the accuracy of solutions. For example, we know that the equilibrium quantities are typically very sensitive to β and that they are less sensitive to d, so it could be that 0.3% deviation in β implies more accuracy decline than 10% deviation in d. Hence, we must have some knowledge of how sensitive the model's variables are to the parameters. Our forward lower error-bound analysis delivers more tractable results.

4 Application 2: a new Keynesian model

We now assess the approximation errors in the conventional new Keynesian model with Calvo-type price frictions and a Taylor (1993) rule; see, e.g., Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2003, 2007), Del Negro et al. (2007). The above literature estimates the model's parameters using data on actual economies, and more recent literature also calibrates and solves such models; e.g., Judd et al. (2011b, 2012), Fernández-Villaverde et al. (2012), Maliar and Maliar (2013), Aruoba and Schorfheide (2013), Gavion et al. (2013); also, see Maliar and Maliar (2014) for a review of numerical method for large scale dynamic economic models. In particular, Judd et al. (2011b, 2012) assess accuracy of non-linear and perturbation methods by constructing residuals in the model's equations and argue that perturbation methods are unreliable for new Keynesian models. In this

section, we perform a direct assessment of accuracy of perturbation methods in the context of new Keynesian models by constructing lower bounds on approximation errors.

4.1 The model

The economy is populated by households, final-good firms, intermediate-good firms, monetary authority and government; see Galí (2008, Chapter 3) for a detailed description of the baseline new Keynesian model.

Households. The representative household solves

$$\max_{\{C_t, L_t, B_t\}_{t=0,\ldots,\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \exp\left(\eta_{u,t}\right) \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \exp\left(\eta_{L,t}\right) \frac{L_t^{1+\vartheta} - 1}{1+\vartheta}\right] \tag{41}$$

$$\text{s.t. } P_t C_t + \frac{B_t}{\exp\left(\eta_{B,t}\right) R_t} + T_t = B_{t-1} + W_t L_t + \Pi_t, \tag{42}$$

where $\left(B_0, \eta_{u,0}, \eta_{L,0}, \eta_{B,0}\right)$ is given; C_t, L_t, and B_t are consumption, labor and nominal bond holdings, respectively; P_t, W_t and R_t are the commodity price, nominal wage and (gross) nominal interest rate, respectively; $\eta_{u,t}$ and $\eta_{L,t}$ are exogenous preference shocks to the overall momentary utility and disutility of labor, respectively; $\eta_{B,t}$ is an exogenous premium in the return to bonds; T_t is lump-sum taxes; Π_t is the profit of intermediate-good firms; $\beta \in (0,1)$ is the discount factor; $\gamma > 0$ and $\vartheta > 0$ are the utility-function parameters. The processes for shocks are

$$\eta_{u,t+1} = \rho_u \eta_{u,t} + \epsilon_{u,t+1}, \qquad \epsilon_{u,t+1} \sim \mathcal{N}\left(0, \sigma_u^2\right), \tag{43}$$

$$\eta_{L,t+1} = \rho_L \eta_{L,t} + \epsilon_{L,t+1}, \qquad \epsilon_{L,t+1} \sim \mathcal{N}\left(0, \sigma_L^2\right), \tag{44}$$

$$\eta_{B,t+1} = \rho_B \eta_{B,t} + \epsilon_{B,t+1}, \qquad \epsilon_{B,t+1} \sim \mathcal{N}\left(0, \sigma_B^2\right), \tag{45}$$

where ρ_u, ρ_L, ρ_B are the autocorrelation coefficients, and σ_u, σ_L, σ_B are the standard deviations of disturbances.

Final-good firms. Perfectly competitive final-good firms produce final goods using intermediate goods. A final-good firm buys $Y_t(i)$ of an intermediate good $i \in [0,1]$ at price $P_t(i)$ and sells Y_t of the final good at price P_t in a perfectly competitive market. The profit-maximization problem is

$$\max_{Y_t(i)} \quad P_t Y_t - \int_0^1 P_t(i) Y_t(i)\, di \tag{46}$$

$$\text{s.t. } Y_t = \left(\int_0^1 Y_t(i)^{\frac{\varepsilon-1}{\varepsilon}}\, di\right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{47}$$

where (47) is a Dixit-Stiglitz aggregator function with $\varepsilon \geq 1$.

Intermediate-good firms. Monopolistic intermediate-good firms produce intermediate goods using labor and are subject to sticky prices. The firm i produces the intermediate good i. To choose labor in each period t, the firm i minimizes the nominal total cost, TC (net of government subsidy v),

$$\min_{L_t(i)} \quad \mathrm{TC}\left(Y_t\left(i\right)\right) = \left(1-v\right) W_t L_t\left(i\right) \tag{48}$$

$$\text{s.t. } Y_t\left(i\right) = \exp\left(\eta_{a,t}\right) L_t\left(i\right), \tag{49}$$

$$\eta_{a,t+1} = \rho_a \eta_{a,t} + \epsilon_{a,t+1}, \qquad \epsilon_{a,t+1} \sim \mathcal{N}\left(0, \sigma_a^2\right), \tag{50}$$

where $L_t\left(i\right)$ is the labor input; $\exp\left(\eta_{a,t}\right)$ is the productivity level; ρ_a is the autocorrelation coefficient; and σ_a is the standard deviation of the disturbance. The firms are subject to Calvo-type price setting: a fraction $1-\theta$ of the firms sets prices optimally, $P_t\left(i\right) = \widetilde{P}_t$, for $i \in [0,1]$, and the fraction θ is not allowed to change the price and maintains the same price as in the previous period, $P_t\left(i\right) = P_{t-1}\left(i\right)$, for $i \in [0,1]$. A reoptimizing firm $i \in [0,1]$ maximizes the current value of the profit over the time when $\widetilde{P}_t$ remains effective,

$$\max_{\widetilde{P}_t} \quad \sum_{j=0}^{\infty} \beta^j \theta^j E_t \left\{ \Lambda_{t+j} \left[\widetilde{P}_t Y_{t+j}\left(i\right) - P_{t+j} \mathrm{mc}_{t+j} Y_{t+j}\left(i\right) \right] \right\} \tag{51}$$

$$\text{s.t. } Y_t\left(i\right) = Y_t \left(\frac{P_t\left(i\right)}{P_t} \right)^{-\varepsilon}, \tag{52}$$

where (52) is the demand for an intermediate good i following from (46), (47); Λ_{t+j} is the Lagrange multiplier on the household's budget constraint (42); mc_{t+j} is the real marginal cost of output at time $t+j$ (which is identical across the firms).

Government. Government finances a stochastic stream of public consumption by levying lump-sum taxes and by issuing nominal debt. The government budget constraint is

$$T_t + \frac{B_t}{\exp\left(\eta_{B,t}\right) R_t} = P_t \frac{\overline{G} Y_t}{\exp\left(\eta_{G,t}\right)} + B_{t-1} + v W_t L_t, \tag{53}$$

where $\frac{\overline{G} Y_t}{\exp\left(\eta_{G,t}\right)} = G_t$ is government spending, $v W_t L_t$ is the subsidy to the intermediate-good firms, and $\eta_{G,t}$ is a government-spending shock,

$$\eta_{G,t+1} = \rho_G \eta_{G,t} + \epsilon_{G,t+1}, \qquad \epsilon_{G,t+1} \sim \mathcal{N}\left(0, \sigma_G^2\right), \tag{54}$$

where ρ_R is the autocorrelation coefficient, and σ_R is the standard deviation of disturbance.

Monetary authority. The monetary authority follows a Taylor rule:

$$R_t = R_* \left(\frac{R_{t-1}}{R_*}\right)^{\mu} \left[\left(\frac{\pi_t}{\pi_*}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_{N,t}}\right)^{\phi_y}\right]^{1-\mu} \exp\left(\eta_{R,t}\right), \tag{55}$$

where R_* is the long-run value of the gross nominal interest rate; π_* is the target inflation; $Y_{N,t}$ is the natural level of output; and $\eta_{R,t}$ is a monetary shock,

$$\eta_{R,t+1} = \rho_R \eta_{R,t} + \epsilon_{R,t+1}, \qquad \epsilon_{R,t+1} \sim \mathcal{N}\left(0, \sigma_R^2\right), \tag{56}$$

where ρ_R is the autocorrelation coefficient, and σ_R is the standard deviation of disturbance.

Natural level of output. The natural level of output $Y_{N,t}$ is the level of output in an otherwise identical economy but without distortions. It is a solution to the following planner's problem

$$\max_{\{C_t, L_t\}_{t=0,\ldots,\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \exp\left(\eta_{u,t}\right) \left[\frac{C_t^{1-\gamma} - 1}{1-\gamma} - \exp\left(\eta_{L,t}\right) \frac{L_t^{1+\vartheta} - 1}{1+\vartheta}\right] \tag{57}$$

$$\text{s.t. } C_t = \exp\left(\eta_{a,t}\right) L_t - G_t, \tag{58}$$

where $G_t = \frac{\overline{G} Y_t}{\exp\left(\eta_{G,t}\right)}$ is given, and $\eta_{u,t+1}$, $\eta_{L,t+1}$, $\eta_{a,t+1}$, and $\eta_{G,t}$ follow the processes (43), (44), (50), and (54), respectively.

4.2 A lower bound on approximation error

We define lower bounds on approximation errors for a new Keynesian model, and we assess such bounds numerically.

4.2.1 Defining a lower error bound

The FOCs of the above model are derived in Appendix B1; see the system of six equations (B36)–(B41). The approximation errors of the equilibrium quantities satisfy the following six equations:

$$\frac{\exp\left(\eta_{u,t} + \eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}} \frac{\left(G_t^{-1} \widehat{C}_t\right)^{1+\vartheta} \left(1 + \delta_{C_t}\right)^{1+\vartheta}}{\left(\widehat{\Delta}_t\right)^{\vartheta} \left(1 + \delta_{\Delta_t}\right)^{\vartheta}}$$

$$+\beta\theta E_t \left\{\widehat{\pi}_{t+1}^{\varepsilon} \left(1 + \delta_{\pi_{t+1}}\right)^{\varepsilon} \widehat{S}_{t+1} \left(1 + \delta_{S_{t+1}}\right)\right\} - \widehat{S}_t \left(1 + \delta_{S_t}\right) = 0, \tag{59}$$

$$\exp\left(\eta_{u,t}\right) G_t^{-1} \widehat{C}_t^{1-\gamma} \left(1+\delta_{C_t}\right)^{1-\gamma}$$
$$+\beta\theta E_t \left\{ \widehat{\pi}_{t+1}^{\varepsilon-1} \left(1+\delta_{\pi_{t+1}}\right)^{\varepsilon-1} \widehat{F}_{t+1} \left(1+\delta_{F_{t+1}}\right) \right\} - \widehat{F}_t \left(1+\delta_{F_t}\right) = 0, \tag{60}$$

$$\frac{\widehat{S}_t \left(1+\delta_{S_t}\right)}{\widehat{F}_t \left(1+\delta_{F_t}\right)} = \left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1} \left(1+\delta_{\pi_t}\right)^{\varepsilon-1}}{1-\theta} \right]^{\frac{1}{1-\varepsilon}}, \tag{61}$$

$$\left[(1-\theta) \left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1} \left(1+\delta_{\pi_t}\right)^{\varepsilon-1}}{1-\theta} \right]^{\frac{\varepsilon}{\varepsilon-1}} + \theta \frac{\widehat{\pi}_t^{\varepsilon} \left(1+\delta_{\pi_t}\right)^{\varepsilon}}{\Delta_{t-1}} \right]^{-1}$$
$$-\widehat{\Delta}_t \left(1+\delta_{\Delta_t}\right) = 0, \tag{62}$$

$$\beta \frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)} \widehat{R}_t \left(1+\delta_{R_t}\right) E_t \left[\frac{\widehat{C}_{t+1}^{-\gamma} \left(1+\delta_{C_{t+1}}\right)^{-\gamma} \exp\left(\eta_{u,t+1}\right)}{\widehat{\pi}_{t+1} \left(1+\delta_{\pi_{t+1}}\right)} \right]$$
$$-\widehat{C}_t^{-\gamma} \left(1+\delta_{C_t}\right)^{-\gamma} = 0, \tag{63}$$

$$R_* \left(\frac{R_{t-1}}{R_*} \right)^{\mu} \left[\left(\frac{\widehat{\pi}_t \left(1+\delta_{\pi_t}\right)}{\pi_*} \right)^{\phi_\pi} \left(\frac{G_t^{-1} \widehat{C}_t \left(1+\delta_{C_t}\right)}{Y_{N,t}} \right)^{\phi_y} \right]^{1-\mu}$$
$$\exp\left(\eta_{R,t}\right) - \widehat{R}_t \left(1+\delta_{R_t}\right) = 0, \tag{64}$$

where hats on the variables denote their approximated values; S_t and F_t are supplementary variables; Δ_t is a measure of price dispersion across firms (see Appendix B1 for details).

We identify the lower error bounds to minimize the least-squares criterion for each t

$$\min_{x_t} \delta_{C_t}^2 + \delta_{F_t}^2 + \delta_{\pi_t}^2 + \delta_{\pi_{t+1}}^2 + \delta_{\Delta_t}^2 + \delta_{S_t}^2 + \delta_{S_{t+1}}^2 + \delta_{F_{t+1}}^2 + \delta_{R_t}^2 + \delta_{C_{t+1}}^2 \tag{65}$$
$$\text{s.t. (59)–(64)},$$

where $x_t \equiv \left\{ \delta_{C_t}, \delta_{F_t}, \delta_{\pi_t}, \delta_{\pi_{t+1}}, \delta_{\Delta_t}, \delta_{S_t}, \delta_{S_{t+1}}, \delta_{F_{t+1}}, \delta_{R_t}, \delta_{C_{t+1}} \right\}$.

4.2.2 Numerical experiments

We set the discount factor at $\beta = 0.99$. To parameterize Taylor rule (55), we use the steady-state interest rate $R_* = \frac{\pi_*}{\beta}$. For the remaining parameters, we consider three alternative parameterizations. The first parameterization mostly corresponds

to the estimated values of the parameters in Smets and Wouters (2003). The second parameterization is mostly taken from the estimates of Del Negro et al. (2007). In the final parameterization, we vary some parameters values relative to the second parameterization; namely, we use the value of $\sigma_L = 1\%$, which is significantly lower than Del Negro's et al. (2007) lower bound on this parameter $\sigma_L = 18.21\%$, and we use the value of target inflation, $\pi_* = 1$ (a zero net target inflation); see Table 4.

Table 4: Alternative parameterizations of the new Keynesian model

	Parameterization 1			Parameterization 2			Parameterization 3		
	min	max	point	min	max	point	min	max	point
γ	0.959	1.902	1.391			1			1
ϑ	1.603	3.481	2.503	0.95	3.19	2.09	0.95	3.19	2.09
ε	2.493	4.236	3.096	3.94	5.16	4.45	3.94	5.16	4.45
ϕ_y	0.037	0.169	0.098	0.03	0.10	0.07	0.03	0.10	0.07
ϕ_π	1.526	1.844	1.688	1.79	2.63	2.21	1.79	2.63	2.21
μ	0.932	0.974	0.956	0.78	0.86	0.82	0.78	0.86	0.82
$\overline{G}$			0.18	0.20	0.26	0.23	0.20	0.26	0.23
θ	0.888	0.922	0.905	0.79	0.87	0.83	0.79	0.87	0.83
π^*			1	1.0461	1.0738	1.0598	1.0461	1.0738	1
ρ_a	0.697	0.910	0.697	0.94	0.97	0.95	0.94	0.97	0.94
ρ_u	0.772	0.894	0.772	0.86	0.97	0.92	0.86	0.97	0.86
ρ_G	0.900	0.977	0.900	0.93	0.97	0.95	0.93	0.97	0.93
ρ_L	0.773	0.952	0.773	0.11	0.37	0.25	0.11	0.37	0.11
σ_a	0.469	0.874	0.469	0.41	0.50	0.45	0.41	0.50	0.41
σ_u	0.237	0.631	0.237	0.36	0.71	0.54	0.36	0.71	0.36
σ_G	0.292	0.385	0.292	0.34	0.42	0.38	0.34	0.42	0.34
σ_L	2.313	5.845	2.313	18.21	64.08	18.21	18.21	64.08	1
σ_R	0.060	0.125	0.090	0.25	0.31	0.28	0.25	0.31	0.25

Notes: Alternative 1 is in line with the estimates of Smets and Wouter (2003), and Alternatives 2 and 3 are in line with the estimates of Del Negro et al. (2006); volatilities are expressed in percentage terms "min" and "max" show the interval of the estimated values of the parameter and "point" is the most plausible value.

We simulate the perturbation solutions using the Dynare's representation of the state space which includes the current endogenous state variables $\{\Delta_{t-1}, R_{t-1}\}$, the past exogenous state variables $\left\{\eta_{u,t-1}, \eta_{L,t-1}, \eta_{B,t-1}, \eta_{a,t-1}, \eta_{R,t-1}, \eta_{G,t-1}\right\}$ and the current disturbances $\{\epsilon_{u,t}, \epsilon_{L,t}, \epsilon_{B,t}, \epsilon_{a,t}, \epsilon_{R,t}, \epsilon_{G,t}\}$. The length of stochastic simulation is 10,200 observations (we eliminate the first 200 observations). To compute the conditional expectation, we use a monomial rule with $2N^2 + 1$ nodes where $N = 6$ is the number of exogenous shocks (see Judd et al., 2011a, for details).

We report the size of approximation errors in Table 5.

The size of approximation errors depends significantly on a specific parameterization. Under Parameterization 3 with a low σ_L, we got a lower bound on approximation errors of order $0.001\% \approx 10^{-2.86}$. However, under the other two parameterizations, the errors are much larger. In particular, under Parameteriza-

Table 5: Approximation errors of the equilibrium allocations in the new Keynesian model

	Errors	δ_{C_t}	δ_{F_t}	δ_{π_t}	$\delta_{\pi_{t+1}}$	δ_{Δ_t}	δ_{S_t}	$\delta_{S_{t+1}}$	$\delta_{F_{t+1}}$	δ_{R_t}	$\delta_{C_{t+1}}$	All
Parameterization 1												
PER1	L_1	-2.47	-2.76	-2.72	-2.49	-2.48	-3.59	-2.64	-2.88	-2.67	-2.70	-2.67
	L_∞	-0.65	-0.63	-1.65	-0.83	-0.60	-1.92	-1.68	-0.97	-1.29	-0.96	-0.60
Parameterization 2												
PER1	L_1	-1.40	-1.56	-1.65	-1.59	-1.59	-1.96	-1.90	-1.48	-1.46	-1.86	-1.61
	L_∞	0.21	0.01	0.00	-0.33	-0.32	-0.66	0.05	0.13	0.11	0.12	0.21
Parameterization 3												
PER1	L_1	-3.01	-3.66	-3.64	-4.19	-3.84	-4.60	-3.45	-3.37	-3.52	-3.90	-3.54
	L_∞	-1.98	-2.25	-2.67	-2.48	-2.52	-2.88	-2.31	-2.17	-2.52	-3.01	-1.98
PER2	L_1	-3.66	-4.47	-3.98	-4.44	-5.19	-4.80	-3.39	-3.83	-3.64	-4.54	-3.91
	L_∞	-2.86	-3.44	-3.50	-3.97	-3.99	-4.46	-2.98	-3.21	-3.46	-3.79	-2.86

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of the lower bounds on approximation errors across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observationsolution (in seconds).

tion 2, the lower error bounds for *PER1* reaches $160\% \approx 10^{0.21}$. Moreover, under Parameterizations 1 and 2, *PER2* solutions are numerically unstable (explosive) in simulation. The fact that errors are so huge even under our optimistic scenario makes these numerical solutions inacceptable for any application!

4.3 Analysis of residuals in model's equations

We consider the following six unit-free residuals in the model's FOCs:

$$\mathcal{R}_t^1(x_t) = \frac{1}{\widehat{S}_t}\left[\frac{\exp\left(\eta_{u,t}+\eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}}\frac{\left(G_t^{-1}\widehat{C}_t\right)^{1+\vartheta}}{\left(\widehat{\Delta}_t\right)^{\vartheta}}+\beta\theta E_t\left\{\widehat{\pi}_{t+1}^{\varepsilon}\widehat{S}_{t+1}\right\}\right]-1, \quad (66)$$

$$\mathcal{R}_t^2(x_t) = \frac{1}{\widehat{F}_t}\left[\exp\left(\eta_{u,t}\right)G_t^{-1}\widehat{C}_t^{1-\gamma}+\beta\theta E_t\left\{\widehat{\pi}_{t+1}^{\varepsilon-1}\widehat{F}_{t+1}\right\}\right]-1, \quad (67)$$

$$\mathcal{R}_t^3(x_t) = \frac{1}{\widehat{\pi}_t}\left[\frac{1}{\theta}+\left(1-\frac{1}{\theta}\right)\left[\frac{\widehat{S}_t}{\widehat{F}_t}\right]^{1-\varepsilon}\right]^{\frac{1}{\varepsilon-1}}-1, \quad (68)$$

$$\mathcal{R}_t^4(x_t) = \frac{1}{\widehat{\Delta}_t}\left[(1-\theta)\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{\varepsilon}{\varepsilon-1}}+\theta\frac{\widehat{\pi}_t^{\varepsilon}}{\Delta_{t-1}}\right]^{-1}-1, \quad (69)$$

$$\mathcal{R}_t^5(x_t) = \frac{1}{\widehat{C}_t}\left\{\frac{\beta\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)}\widehat{R}_t E_t\left[\frac{\widehat{C}_{t+1}^{-\gamma}\exp\left(\eta_{u,t+1}\right)}{\widehat{\pi}_{t+1}}\right]\right\}^{-1/\gamma}-1, \quad (70)$$

$$\mathcal{R}_t^6(x_t) = \frac{R_*}{\widehat{R}_t}\left(\frac{R_{t-1}}{R_*}\right)^{\mu}\left[\left(\frac{\widehat{\pi}_t}{\pi_*}\right)^{\phi_\pi}\left(\frac{G_t^{-1}\widehat{C}_t}{Y_{N,t}}\right)^{\phi_y}\right]^{1-\mu}\exp\left(\eta_{R,t}\right)-1. \quad (71)$$

We report the residuals in Table 6. The accuracy implications from residuals

Table 6: Residuals of the equilibrium equations in the new Keynesian model

	Residuals	R^1	R^2	R^3	R^4	R^5	R^6	*All*
Parameterization 1								
PER1	L_1	-2.00	-2.76	-2.80	-2.39	-2.61	-4.81	-2.48
	L_∞	-0.48	-1.62	-1.99	-0.36	-1.42	-3.57	-0.36
Parameterization 2								
PER1	L_1	-1.77	-1.35	0.07	-1.69	-2.16	-3.40	-0.67
	L_∞	-0.24	2.13	3.92	-0.12	-0.87	-1.35	3.92
Parameterization 3								
PER1	L_1	-3.27	-3.71	-3.73	-3.68	-3.85	-5.22	-3.67
	L_∞	-2.29	-2.87	-2.72	-2.41	-2.68	-4.14	-2.29
PER2	L_1	-3.42	-4.15	-3.82	-4.70	-5.34	-6.66	-3.98
	L_∞	-2.90	-3.55	-3.66	-3.42	-4.07	-5.05	-2.90

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of residuals in the model's equations errors across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observations.

in the model's equations are similar to those from the approximation errors. The maximum residuals are quite low, $0.1\% \approx 10^{-2.29}$, for a *PER2* solution obtained under Parameterization 3, however, they are enormous, $800000\% \approx 10^{3.92}$, for a *PER1* solution obtained under Parameterization 2 with a high σ_L. For all parameterizations, the maximum residuals in Table 6 are substantially larger than lower bounds on approximation errors in Table 5.

As in the optimal growth model, for the new Keynesian model, we define the residuals in model's equations (66)-(71) in a way that is comparable to approximation errors in the corresponding model's variables (59)-(64). Again, in the case of the residuals, we compute an approximation error in each variable by assuming that approximation errors in all other variables are zeros which is mutually inconsistent. In turn, lower bounds are constructed in the way that is consistent with all model's equations and that minimizes approximation errors in all the variables. The difference between the residuals and lower error bounds is very large in some of our experiments. Our results suggest that in the presence of strong nonlinearities, the analysis of residuals and lower error-bound analysis may lead to qualitatively different inferences about the accuracy of numerical solutions.

5 Conclusion

The conventional upper error-bound analysis focuses on worst case scenarios and provides sufficient conditions for accuracy of numerical solutions. In this paper, we introduce a complementary lower error-bound analysis that focuses on certain best-case scenarios and provides a necessary condition for accuracy of numerical solutions. We specifically construct the smallest possible (optimistic) approximation errors that are consistent with some subset of model's equations. Even if these optimistic errors are too large, we conclude that a numerical solution is inaccurate.

A potential shortcoming of our test is that it may fail to reject inaccurate solutions because some inaccurate solutions may appear to be sufficiently accurate under best-case scenarios. But one of the two studied models - a stylized new Keynesian model - failed to pass even this relatively undemanding test under some empirically relevant parameterizations. Upper error bounds are unknown for new Keynesian models but they are also unnecessary in those cases when an approximate solution fails to satisfy even necessary conditions for accuracy. Thus, our simple accuracy test is powerful enough to detect and to discard inaccurate solutions in practically relevant applications.

References

[1] Adjemian, S., H. Bastani, M. Juillard, F. Mihoubi, G. Perendia, M. Ratto and S. Villemot, (2011). Dynare: Reference Manual, Version 4. Dynare Working Papers 1, CEPREMAP.

[2] Aruoba, S., J. Fernández-Villaverde and J. Rubio-Ramírez, (2006). Comparing solution methods for dynamic equilibrium economies. Journal of Economic Dynamics and Control 30, 2477-2508.

[3] Aruoba, S. and F. Schorfheide, (2013). Macroeconomic dynamics near ZLB: a tale of two equilibria. NBER working paper 19248.

[4] Bertsekas, D. and J. Tsitsiklis, (1996). Neuro-Dynamic Programming, Bellmont, Massachusetts: Athena Scientific.

[5] Christiano, L., M. Eichenbaum and C. Evans, (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy 113/1, 1-45.

[6] Christiano, L. and D. Fisher, (2000). Algorithms for solving dynamic models with occasionally binding constraints. Journal of Economic Dynamics and Control 24, 1179-1232.

[7] Collard, F., and M. Juillard, (2001). Accuracy of stochastic perturbation methods: the case of asset pricing models, Journal of Economic Dynamics and Control, 25, 979-999.

[8] Del Negro, M., F. Schorfheide, F. Smets and R. Wouters, (2007). On the fit of new Keynesian models. Journal of Business and Economic Statistics 25 (2), 123-143.

[9] Den Haan, W. and A. Marcet, (1990). Solving the stochastic growth model by parameterized expectations. Journal of Business and Economic Statistics 8, 31-34.

[10] Den Haan, W., (2010), Comparison of solutions to the incomplete markets model with aggregate uncertainty. Journal of Economic Dynamics and Control 34, 4–27.

[11] Den Haan, W. and A. Marcet, (1994). Accuracy in Simulations. Review of Economic Studies 61, 3-17.

[12] Fernández-Villaverde, J., G. Gordon, P. Guerrón-Quintana, and J. Rubio-Ramírez, (2012). Nonlinear adventures at the zero lower bound. NBER working paper 18058.

[13] Galí, J., (2008). Monetary Policy, Inflation and the Business Cycles: An Introduction to the New Keynesian Framework. Princeton University Press: Princeton, New Jersey.

[14] Gavion, W.T., B.D.Keen, A.W.Richter, and N.A. Throckmorton, (2013). Global dynamics at the zero lower bound, Federal Reserve Board St. Louis working paper 2013-007B.

[15] Gaspar, J. and K. Judd, (1997). Solving large-scale rational-expectations models. Macroeconomic Dynamics 1, 45-75.

[16] Higham, N., (1996). Accuracy and Stability of Numerical Algorithms. Philadelphia: SIAM.

[17] Judd, K., (1992). Projection methods for solving aggregate growth models. Journal of Economic Theory 58, 410-452.

[18] Judd, K., (1998). Numerical Methods in Economics. Cambridge, MA: MIT Press.

[19] Judd, K. and S. Guu, (1993). Perturbation solution methods for economic growth models, in: H. Varian, (Eds.), Economic and Financial Modeling with Mathematica, Springer Verlag, pp. 80-103.

[20] Judd, K., L. Maliar and S. Maliar, (2011). Numerically stable and accurate stochastic simulation approaches for solving dynamic models. Quantitative Economics 2, 173-210.

[21] Judd, K., L. Maliar and S. Maliar, (2011b). A cluster grid projection algorithm: solving problems with high dimensionality. Manuscript available at http://stanford.edu/~maliarl.

[22] Judd, K., L. Maliar and S. Maliar, (2012). Merging simulation and projection approaches to solve high-dimensional problems. NBER working paper 18501.

[23] Juillard, M. and S. Villemot, (2011). Multi-country real business cycle models: Accuracy tests and testing bench. Journal of Economic Dynamics and Control 35, 178–185.

[24] Kogan, L. and I. Mitra, (2013). Accuracy Verification for Numerical Solutions of Equilibrium Models. Manuscript http://web.mit.edu/lkogan2/www/ResearchFrameset.htm.

[25] Kollmann, R., (2002). Monetary policy rules in the open economy: effects on welfare and business cycles. Journal of Monetary Economics 49, 989-1015.

[26] Kollmann, R., S. Kim and J. Kim, (2011a). Solving the multi-country real business cycle model using a perturbation method. Journal of Economic Dynamics and Control 35, 203-206.

[27] Kollmann, R., S. Maliar, B. Malin and P. Pichler, (2011). Comparison of solutions to the multi-country real business cycle model. Journal of Economic Dynamics and Control 35, 186-202.

[28] Krueger, D. and F. Kubler, (2004). Computing equilibrium in OLG models with production. Journal of Economic Dynamics and Control 28, 1411-1436.

[29] Kubler, F. and K. Schmedders, (2005). Approximate versus exact equilibria in dynamic economies. Econometrica 73, 1205-1235.

[30] Maliar, L. and S. Maliar, (2013). Merging Simulation and Projection Approaches to Solve High-Dimensional Problems with an Application to a New Keynesian model. Quantitative Economics (forthcoming).

[31] Maliar, L. and S. Maliar, (2014). Numerical methods for large scale dynamic economic models. In: Schmedders, K., Judd, K. (Eds.), Handbook of Computational Economics, vol. 3. , Elsevier Science, Amsterdam.

[32] Maliar, S., L. Maliar and K. Judd, (2011). Solving the multi-country real business cycle model using ergodic set methods. Journal of Economic Dynamic and Control 35, 207–228.

[33] National Research Council. (2012). Assessing the reliability of complex models: mathematical and statistical foundations of verification, validation, and uncertainty quantification. Washington D.C.: The National Academies Press.

[34] Oberkamf, W. and C. Roy, (2010). Verification and Validation in Scientific Computing, Cambridge: Cambridge University Press.

[35] Peralta-Alva, A. and M. Santos, (2014). Analysis of numerical errors, in: K. Schmedders and K. Judd (Eds), Handbook of Computational Economics, Vol. 3, Amsterdam: Elsevier Science, pp. 517-556.

[36] Santos, M., (1999). Numerical solution of dynamic economic models, in: J. Taylor and M. Woodford (Eds.), Handbook of Macroeconomics, Amsterdam: Elsevier Science, pp. 312-382.

[37] Santos, M., (2000). Accuracy of numerical solutions using the Euler equations residuals. Econometrica 68, 1377-1402.

[38] Santos, M. and A. Peralta-Alva, (2005), Accuracy of simulations for stochastic dynamic models. Econometrica 73, 1939-1976.

[39] Santos, M. and J. Vigo-Aguiar, (1998). Analysis of a numerical dynamic programming algorithm applied to economic models. Econometrica 66, 409-426.

[40] Sims, C. (1990). Solving the stochastic growth model by backsolving with a particular nonlinear form for the decision rule. Journal of Business & Economic Statistics 8(1), 45-47.

[41] Smets, F. and R. Wouters, (2003). An estimated dynamic stochastic general equilibrium model of the Euro area. Journal of the European Economic Association 1(5), 1123-1175.

[42] Smets, F. and R. Wouters, (2007). Shocks and frictions in US business cycles: a Bayesian DSGE approach. American Economic Review 97 (3), 586-606.

[43] Sundarama, R., (2006). A First Course in Optimization Theory. Cambridge: Cambridge University Press.

[44] Taylor, J. and H. Uhlig, (1990). Solving nonlinear stochastic growth models: a comparison of alternative solution methods. Journal of Business and Economic Statistics 8, 1-17.

[45] Wilkinson, J., (1963). Rounding Errors in Algebraic Processes. Englewood Cliffs, NJ: Prentice-Hall.

Supplement to "Lower Bounds on Approximation Errors: Testing the Hypothesis That a Numerical Solution Is Accurate": Appendices

Kenneth L. Judd
Lilia Maliar
Serguei Maliar

In Appendices A and B, we describe details of numerical experiments for the neoclassical stochastic growth model and the new Keynesian model studied in Sections 3 and 4, respectively.

Appedix A: Neoclassical stochastic growth model

In this section, we describe the construction of a lower bound on approximation errors in a neoclassical stochastic growth model. We first construct a rough approximation for such a bound using linearized model's equations. We then construct a lower bound more accurately using nonlinear model's equations. We finally describe additional experiments in which we weigh approximation errors in different variables by different weights.

A1. Constructing approximation errors using linearized model's equations

Euler equation. Let us assume a CRRA utility function $u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$. Under this utility function, Euler equation (33) is

$$\widehat{c}_t^{-\gamma}(1+\delta_{c_t})^{-\gamma} - \beta E_t\left\{\widehat{c}_{t+1}^{-\gamma}\left(1+\delta_{c_{t+1}}\right)^{-\gamma} \cdot \left[1-d+\alpha\exp(\theta_{t+1})A\widehat{k}_{t+1}^{\alpha-1}\left(1+\delta_{k_{t+1}}\right)^{\alpha-1}\right]\right\} = 0. \quad \text{(A1)}$$

Finding a first-order Taylor expansion of (A1) around $\delta_{c_t} \to 0$, $\delta_{c_{t+1}} \to 0$, $\delta_{k_{t+1}} \to 0$ (in particular, using $(1+x)^{\alpha} \simeq 1+\alpha x$) and omitting a second-order term

$\delta_{c_{t+1}}\delta_{k_{t+1}} \approx 0$, we have

$$\widehat{c_t}^{-\gamma} - \gamma\delta_{c_t}\widehat{c_t}^{-\gamma} - \beta E_t\left\{\widehat{c_{t+1}}^{-\gamma}\left(1-d+\alpha\exp\left(\theta_{t+1}\right)A\widehat{k}_{t+1}^{\alpha-1}\right)\right\}$$
$$+\beta E_t\left\{\widehat{c_{t+1}}^{-\gamma}\gamma\delta_{c_{t+1}}\left(1-d+\alpha\exp\left(\theta_{t+1}\right)A\widehat{k}_{t+1}^{\alpha-1}\right)\right\}$$
$$-\beta E_t\left\{\widehat{c_{t+1}}^{-\gamma}\left(\alpha\exp\left(\theta_{t+1}\right)A\widehat{k}_{t+1}^{\alpha-1}\left(\alpha-1\right)\delta_{k_{t+1}}\right)\right\} = 0.$$

In terms of $\delta_{c_{t+1}}$, the previous equation can be written as

$$1 - \gamma\delta_{c_t} - X_{1t} + \gamma\delta_{c_{t+1}}X_{1t} - \left(\alpha-1\right)\delta_{k_{t+1}}X_{2t} = 0,$$

where

$$X_{1t} \equiv \beta E_t\left\{\frac{\widehat{c_{t+1}}^{-\gamma}}{\widehat{c_t}^{-\gamma}}\left(1-d+\alpha\exp\left(\theta_{t+1}\right)A\widehat{k}_{t+1}^{\alpha-1}\right)\right\},$$
$$X_{2t} \equiv \beta E_t\left\{\frac{\widehat{c_{t+1}}^{-\gamma}}{\widehat{c_t}^{-\gamma}}\left(\alpha\exp\left(\theta_{t+1}\right)A\widehat{k}_{t+1}^{\alpha-1}\right)\right\}.$$

Combining the terms yields a linear equation in δ's,

$$a_t^{11}\delta_{c_t} + a_t^{12}\delta_{k_{t+1}} + a_t^{13}\delta_{c_{t+1}} = b_t^1, \tag{A2}$$

where $a_t^{11} \equiv -\gamma$, $a_t^{12} \equiv -\left(\alpha-1\right)X_{2t}$, $a_t^{13} \equiv \gamma X_{1t}$, and $b_t^1 \equiv X_{1t} - 1$. Therefore, we obtain an equation for $\delta_{c_{t+1}}$ in terms of δ_{c_t} and $\delta_{k_{t+1}}$

$$\delta_{c_{t+1}} = \frac{b_t^1}{a_t^{13}} - \frac{a_t^{11}}{a_t^{13}}\delta_{c_t} - \frac{a_t^{12}}{a_t^{13}}\delta_{k_{t+1}}. \tag{A3}$$

Budget constraint. We rewrite budget constraint (32) as

$$\widehat{c_t} + \delta_{c_t}\widehat{c_t} + \widehat{k}_{t+1} + \delta_{k_{t+1}}\widehat{k}_{t+1} - \left(1-d\right)k_t - \exp\left(\theta_t\right)Ak_t^{\alpha} = 0. \tag{A4}$$

Thus, we have

$$a_t^{21}\delta_{c_t} + a_t^{22}\delta_{k_{t+1}} = b_t^{21},$$

where $a_t^{21} \equiv \widehat{c_t}$, $a_t^{22} \equiv \widehat{k}_{t+1}$, and $b_t^{21} \equiv \left(1-d\right)k_t + \exp\left(\theta_t\right)Ak_t^{\alpha} - \widehat{c_t} - \widehat{k}_{t+1}$, so we write

$$\delta_{k_{t+1}} = \frac{b_t^{21}}{a_t^{22}} - \frac{a_t^{21}}{a_t^{22}}\delta_{c_t}. \tag{A5}$$

Note that substituting $\delta_{k_{t+1}}$ from (A5) into (A2), we obtain a relation between $\delta_{c_{t+1}}$ and δ_{c_t}

$$\delta_{c_{t+1}} = \frac{1}{a_t^{13}}\left(b_t^1 - \frac{a_t^{12}}{a_t^{22}}b_2\right) + \frac{1}{a_t^{13}}\left(\frac{a_t^{12}a_t^{21}}{a_t^{22}} - a_t^{11}\right)\delta_{c_t}. \tag{A6}$$

Minimization problem. The minimization problem (34) in period t is given by

$$\min_{\delta_{c_t},\delta_{k_{t+1}},\delta_{c_{t+1}}} \delta_{c_t}^2 + \delta_{k_{t+1}}^2 + \delta_{c_{t+1}}^2 \quad \text{s.t. (A5), (A6).} \tag{A7}$$

There are two possible ways of solving (A7). One option is to find FOC of (A7) analytically and to solve for a minimum in a closed form and the other option is to solve (A7) numerically.

To implement the first option, we find the FOC with respect to δ_{c_t},

$$2\delta_{c_t} + 2\delta_{k_{t+1}} \frac{d\delta_{k_{t+1}}}{d\delta_{c_t}} + 2\delta_{c_{t+1}} \frac{d\delta_{c_{t+1}}}{d\delta_{c_t}} = 0.$$

After deriving $\frac{d\delta_{c_{t+1}}}{d\delta_{c_t}}$ and $\frac{d\delta_{k_{t+1}}}{d\delta_{c_t}}$ from (A5), (A6) and substituting them into the above FOC, we obtain

$$\delta_{c_t} = -Q_t/W_t,$$

where

$$\begin{aligned} Q_t &= \frac{1}{\left(a_t^{13}\right)^2}\left[\frac{a_t^{12}a_t^{21}}{a_t^{22}} - a_t^{11}\right]\left[b_1 - \frac{a_{12}}{a_{22}}b_2\right] - b_2\frac{a_{21}}{\left(a_{22}\right)^2}, \\ W_t &= 1 + \frac{1}{\left(a_t^{13}\right)^2}\left[\frac{a_t^{12}a_t^{21}}{a_t^{22}} - a_t^{11}\right]^2 + \left(\frac{a_t^{21}}{a_t^{22}}\right)^2. \end{aligned}$$

Thus, we get an explicit expression for δ_{c_t}. The approximation errors $\delta_{c_{t+1}}$ and $\delta_{k_{t+1}}$ are determined by (A5) and (A6).

To implement the second option, i.e., to solve problem (A7) numerically, we use a quadratic programming routine (specifically, "quadprog" in MATLAB) . We use approximation errors which we obtain from linearizing conditions as an initial guess for a nonlinear procedure, discussed in Appendix A2.

A2. Constructing approximation errors using nonlinear model's equations

Budget constraint (A4) yields

$$\delta_{k_{t+1}} = \frac{(1-d)\,k_t + \exp(\theta_t)\,Ak_t^{\alpha} - \widehat{c}_t\,(1+\delta_{c_t})}{\widehat{k}_{t+1}} - 1. \tag{A8}$$

From budget constraint (A4), we also get

$$\left[\widehat{k}_{t+1}\left(1+\delta_{k_{t+1}}\right)\right]^{\alpha-1} = \left[(1-d)\,k_t + \exp(\theta_t)\,Ak_t^{\alpha} - \widehat{c}_t\,(1+\delta_{c_t})\right]^{\alpha-1}.$$

Substituting the latter equation into Euler equation (A1), we have

$$\widehat{c}_t^{-\gamma}\left(1+\delta_{c_t}\right)^{-\gamma} = \beta E_t\left\{\widehat{c}_{t+1}^{-\gamma}\left(1+\delta_{c_{t+1}}\right)^{-\gamma}(1-d)\right\} + \left[(1-d)\,k_t + \exp\left(\theta_t\right)Ak_t^{\alpha} - \widehat{c}_t\left(1+\delta_{c_t}\right)\right]^{\alpha-1} \cdot \beta E_t\left\{\widehat{c}_{t+1}^{-\gamma}\left(1+\delta_{c_{t+1}}\right)^{-\gamma}\alpha\exp\left(\theta_{t+1}\right)A\right\}.$$

Expressing $\delta_{c_{t+1}}$ yields

$$\delta_{c_{t+1}} = \widehat{c}_t\left(1+\delta_{c_t}\right)\cdot\beta E_t\left\{\widehat{c}_{t+1}^{-\gamma}(1-d)\right\} + \left[(1-d)\,k_t + \exp\left(\theta_t\right)Ak_t^{\alpha} - \widehat{c}_t\left(1+\delta_{c_t}\right)\right]^{\alpha-1} \times \beta E_t\left\{\widehat{c}_{t+1}^{-\gamma}\alpha\exp\left(\theta_{t+1}\right)A\right\}^{1/\gamma} - 1. \quad \text{(A9)}$$

Therefore, problem (34) becomes

$$\min_{\delta_{c_t},\delta_{k_{t+1}},\delta_{c_{t+1}}} \delta_{c_t}^2 + \delta_{k_{t+1}}^2 + \delta_{c_{t+1}}^2 \quad \text{s.t.(A9), (A8).} \qquad \text{(A10)}$$

This representation is more convenient relative to (34) because the resulting minimization problem contains just one unknown, δ_{c_t}. To solve problem (A10), we use a nonlinear optimization routine (specifically, "fminsearch" in MATLAB).

A3. Weighting approximation errors in different variables by their relative importance

As is argued in the main text, we can construct a lower bound on approximation errors to minimize the weighted sum of squared approximation errors (36). By changing weights, we can vary the relative importance attached to the errors in different variables. Within this general framework, we can consider three limiting sets of weights (i) $w_1 \to \infty$, $w_2 = w_3$, (ii) $w_2 \to \infty$, $w_1 = w_3$, and (iii) $w_3 \to \infty$, $w_1 = w_2$ in which cases, respectively, δ_{c_t}, $\delta_{k_{t+1}}$ and $\delta_{c_{t+1}}$ are forced to be zero and the remaining two errors are minimized according to the least-squares criterion.

How can this modification effect the lower bound? Clearly, the resulting lower bound can increase but not decrease. To see the point observe that in the baseline case (34), we split approximation errors between three variables in the way which is most favorable for accuracy. Now, we split approximation errors between two variables forcing the remaining error to be equal to zero. As a result, the size of the approximation errors will increase in the two variables that are assumed to be computed with errors. For example, case (iii) would make sense if we believe that we have obtained a true distribution of future consumption c_{t+1}, and then all the

Table 7: Approximation errors under the assumption of zero approximation errors in current consumption

	$\gamma = \frac{1}{10}$			$\gamma = 1$			$\gamma = 10$		
Errors	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$
PER1									
L_1	$-\infty$	-4.12	-3.64	$-\infty$	-4.12	-4.42	$-\infty$	-3.74	-4.09
L_∞	$-\infty$	-3.02	-2.59	$-\infty$	-3.04	-3.58	$-\infty$	-2.61	-3.51
PER2									
L_1	$-\infty$	-5.80	-5.32	$-\infty$	-5.69	-5.98	$-\infty$	-4.75	-5.30
L_∞	$-\infty$	-4.41	-4.22	$-\infty$	-4.44	-4.76	$-\infty$	-3.65	-4.05

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of the lower bounds on approximation errors across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observations; and γ is the coefficient of risk aversion.

Table 8: Approximation errors under the assumption of zero approximation errors in capital

	$\gamma = \frac{1}{10}$			$\gamma = 1$			$\gamma = 10$		
Errors	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$
PER1									
L_1	-3.03	$-\infty$	-2.93	-3.03	$-\infty$	-3.02	-2.69	$-\infty$	-2.70
L_∞	-1.94	$-\infty$	-1.85	-1.95	$-\infty$	-1.94	-1.65	$-\infty$	-1.66
PER2									
L_1	-4.71	$-\infty$	-4.61	-6.30	$-\infty$	-4.59	-3.66	$-\infty$	-3.66
L_∞	-3.33	$-\infty$	-3.27	-3.35	$-\infty$	-3.33	-2.65	$-\infty$	-2.66

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of the lower bounds on approximation errors across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observations; and γ is the coefficient of risk aversion.

Table 9: Approximation errors under the assumption of zero approximation errors in future consumption

	$\gamma = \frac{1}{10}$			$\gamma = 1$			$\gamma = 10$		
Errors	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$	δ_{c_t}	$\delta_{k_{t+1}}$	$\delta_{c_{t+1}}$
PER1									
L_1	-3.65	-4.03	$-\infty$	-4.42	-4.11	$-\infty$	-4.09	-3.75	$-\infty$
L_∞	-2.60	-2.94	$-\infty$	-3.58	-3.03	$-\infty$	-3.51	-2.62	$-\infty$
PER2									
L_1	-5.33	-5.71	$-\infty$	-5.98	-5.68	$-\infty$	-5.30	-4.74	$-\infty$
L_∞	-4.23	-4.37	$-\infty$	-4.76	-4.43	$-\infty$	-4.05	-3.65	$-\infty$

[a] Notes: PER1 and PER2 denote the first- and second-order perturbation solutions; L_1 and L_∞ are, repectively, the average and maximum of absolute values of the lower bounds on approximation errors across optimality condition and test points (in log10 units) on a stochastic simulation of 10,000 observations; and γ is the coefficient of risk aversion.

errors are absorbed by c_t and k_{t+1}. The results for cases (i), (ii), (iii) distinguished above are provided in Tables 7, 8 and 9, respectively.

The comparison of the three tables shows that the lower error bound is the largest when we set $\delta_{k_{t+1}} = 0$, namely, it is $10^{-1.65} = 2\%$ and $10^{-2.65} = 0.2\%$ for *PER1* and *PER2*, respectively. As is seen from Tables 7 and 9, when $\delta_{c_t} = 0$ and $\delta_{c_{t+1}} = 0$, the results are almost identical to those in Table 1. For example, when $\delta_{c_{t+1}} = 0$, the lower bounds on the size of the errors in the model's variables are $10^{-2.6} = 0.25\%$ and $10^{-3.65} = 0.023\%$ for *PER1* and *PER2*, respectively.

Appendix B: New Keynesian model

In this appendix, we derive FOCs of the new Keynesian model outlined in Section 4, and we describe the details of our error bound analysis. Again, we first construct a rough approximation of the lower bound on approximation errors using linearized model's equations, and we then construct such a bound accurately using nonlinear model's equations.

B1. Deriving FOCs

Households. The FOCs of the household's problem (41)–(45) with respect to C_t, L_t and B_t are

$$\Lambda_t = \frac{\exp\left(\eta_{u,t}\right) C_t^{-\gamma}}{P_t}, \tag{B1}$$

$$\exp\left(\eta_{u,t} + \eta_{L,t}\right) L_t^{\vartheta} = \Lambda_t W_t, \tag{B2}$$

$$\exp\left(\eta_{u,t}\right) C_t^{-\gamma} = \beta \exp\left(\eta_{B,t}\right) R_t E_t \left[\frac{\exp\left(\eta_{u,t+1}\right) C_{t+1}^{-\gamma}}{\pi_{t+1}}\right], \tag{B3}$$

where Λ_t is the Lagrange multiplier associated with the household's budget constraint (42). After combining (B1) and (B2), we get

$$\exp\left(\eta_{L,t}\right) L_t^{\vartheta} C_t^{\gamma} = \frac{W_t}{P_t}. \tag{B4}$$

Final-good producers. The FOC of the final-good producer's problem (46), (47) with respect to $Y_t(i)$ yields the demand for the ith intermediate good

$$Y_t(i) = Y_t \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon}. \tag{B5}$$

Substituting the condition (B5) into (47), we obtain

$$P_t = \left(\int_0^1 P_t(i)^{1-\varepsilon} \, di\right)^{\frac{1}{1-\varepsilon}}. \tag{B6}$$

Intermediate-good producers. The FOC of the cost-minimization problem (48)–(50) with respect to $L_t(i)$ is

$$\Theta_t = \frac{(1-v) W_t}{\exp\left(\eta_{a,t}\right)}, \tag{B7}$$

where Θ_t is the Lagrange multiplier associated with (49). The derivative of the total cost in (48) is the nominal marginal cost, $\mathrm{MC}_t(i)$,

$$\mathrm{MC}_t(i) \equiv \frac{d\mathrm{TC}\left(Y_t(i)\right)}{dY_t(i)} = \Theta_t. \tag{B8}$$

The conditions (B7) and (B8) taken together imply that the real marginal cost is the same for all firms,

$$\mathrm{mc}_t(i) = \frac{(1-v)}{\exp\left(\eta_t^a\right)} \cdot \frac{W_t}{P_t} = \mathrm{mc}_t. \tag{B9}$$

The FOC of the reoptimizing intermediate-good firm with respect to $\widetilde{P}_t$ is

$$E_t \sum_{j=0}^{\infty} (\beta\theta)^j \Lambda_{t+j} Y_{t+j} P_{t+j}^{\varepsilon+1} \left[\frac{\widetilde{P}_t}{P_{t+j}} - \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+j} \right] = 0 \tag{B10}$$

From the household's FOC (B1), we have

$$\Lambda_{t+j} = \frac{\exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma}}{P_{t+j}}. \tag{B11}$$

Substituting (B11) into (B10), we get

$$E_t \sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma} Y_{t+j} P_{t+j}^{\varepsilon} \left[\frac{\widetilde{P}_t}{P_{t+j}} - \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+j} \right] = 0. \tag{B12}$$

Let us define $\chi_{t,j}$ such that

$$\chi_{t,j} \equiv \begin{cases} 1 \text{ if } j = 0 \\ \frac{1}{\pi_{t+j} \cdot \pi_{t+j-1} \cdots \pi_{t+1}} \text{ if } j \geq 1 \end{cases} . \tag{B13}$$

Then $\chi_{t,j} = \chi_{t+1,j-1} \cdot \frac{1}{\pi_{t+1}}$ for $j > 0$. Therefore, (B12) becomes

$$E_t \sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma} Y_{t+j} \chi_{t,j}^{-\varepsilon} \left[\widetilde{\widetilde{p}}_t \chi_{t,j} - \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+j} \right] = 0, \tag{B14}$$

where $\widetilde{\widetilde{p}}_t \equiv \frac{\widetilde{P}_t}{P_t}$. We express $\widetilde{\widetilde{p}}_t$ from (B14) as follows

$$\widetilde{\widetilde{p}}_t = \frac{E_t \sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma} Y_{t+j} \chi_{t,j}^{-\varepsilon} \frac{\varepsilon}{\varepsilon-1} \mathrm{mc}_{t+j}}{E_t \sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma} Y_{t+j} \chi_{t,j}^{1-\varepsilon}} \equiv \frac{S_t}{F_t}. \tag{B15}$$

Let us find recursive representations for S_t and F_t. For S_t, we have

$$\begin{aligned}
S_t &\equiv E_t \sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma} Y_{t+j} \chi_{t,j}^{-\varepsilon} \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+j} \\
&= \frac{\varepsilon}{\varepsilon - 1} \exp\left(\eta_{u,t}\right) C_t^{-\gamma} Y_t \mathrm{mc}_t \\
&+ \beta\theta E_t \left\{ \sum_{j=1}^{\infty} (\beta\theta)^{j-1} \exp\left(\eta_{u,t+j}\right) C_{t+j}^{-\gamma} Y_{t+j} \left(\frac{\chi_{t+1,j-1}}{\pi_{t+1}} \right)^{-\varepsilon} \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+j} \right\} \\
&= \frac{\varepsilon}{\varepsilon - 1} \exp\left(\eta_{u,t}\right) C_t^{-\gamma} Y_t \mathrm{mc}_t \\
&+ \beta\theta E_t \left\{ \frac{1}{\pi_{t+1}^{-\varepsilon}} \sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+1+j}\right) C_{t+1+j}^{-\gamma} Y_{t+1+j} \chi_{t+1,j}^{-\varepsilon} \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+1+j} \right\} \\
&= \frac{\varepsilon}{\varepsilon - 1} \exp\left(\eta_{u,t}\right) C_t^{-\gamma} Y_t \mathrm{mc}_t \\
&+ \beta\theta E_t \left\{ \frac{1}{\pi_{t+1}^{-\varepsilon}} E_{t+1} \left(\sum_{j=0}^{\infty} (\beta\theta)^j \exp\left(\eta_{u,t+1+j}\right) C_{t+1+j}^{-\gamma} Y_{t+1+j} \chi_{t+1,j}^{-\varepsilon} \frac{\varepsilon}{\varepsilon - 1} \mathrm{mc}_{t+1+j} \right) \right\} \\
&= \frac{\varepsilon}{\varepsilon - 1} \exp\left(\eta_{u,t}\right) C_t^{-\gamma} Y_t \mathrm{mc}_t + \beta\theta E_t \left\{ \pi_{t+1}^{\varepsilon} S_{t+1} \right\}.
\end{aligned}$$

Substituting mc_t from $(B9)$ into the above recursive formula for S_t, we have

$$S_t = \frac{\varepsilon}{\varepsilon - 1} \exp\left(\eta_{u,t}\right) C_t^{-\gamma} Y_t \frac{(1 - v)}{\exp\left(\eta_{a,t}\right)} \cdot \frac{W_t}{P_t} + \beta\theta E_t \left\{ \pi_{t+1}^{\varepsilon} S_{t+1} \right\}. \tag{C16}$$

Substituting $\frac{W_t}{P_t}$ from (B4) into (C16), we get

$$S_t = \frac{\varepsilon}{\varepsilon - 1} \exp\left(\eta_{u,t}\right) Y_t \frac{(1 - v)}{\exp\left(\eta_{a,t}\right)} \cdot \exp\left(\eta_{L,t}\right) L_t^{\vartheta} + \beta\theta E_t \left\{ \pi_{t+1}^{\varepsilon} S_{t+1} \right\}. \tag{B16}$$

For F_t, the corresponding recursive formula is

$$F_t = \exp\left(\eta_{u,t}\right) C_t^{-\gamma} Y_t + \beta\theta E_t \left\{ \pi_{t+1}^{\varepsilon - 1} F_{t+1} \right\}. \tag{B17}$$

Aggregate price relationship. The condition (B6) can be rewritten as

$$\begin{aligned}
P_t &= \left(\int_0^1 P_t(i)^{1-\varepsilon} \, di \right)^{\frac{1}{1-\varepsilon}} = \\
&\left[\int_{\text{reopt.}} P_t(i)^{1-\varepsilon} \, di + \int_{\text{non-reopt.}} P_t(i)^{1-\varepsilon} \, di \right]^{\frac{1}{1-\varepsilon}},
\end{aligned} \tag{B18}$$

where "reopt." and "non-reopt." denote, respectively, the firms that reoptimize and do not reoptimize their prices at t.

Note that $\int_{\text{non-reopt.}} P_t(i)^{1-\varepsilon} di = \int_0^1 P(j)^{1-\varepsilon} \omega_{t-1,t}(j) dj$, where $\omega_{t-1,t}(j)$ is the measure of non-reoptimizers at t that had the price $P(j)$ at $t-1$. Furthermore, $\omega_{t-1,t}(j) = \theta\omega_{t-1}(j)$, where $\omega_{t-1}(j)$ is the measure of firms with the price $P(j)$ in $t-1$, which implies

$$\int_{\text{non-reopt.}} P_t(i)^{1-\varepsilon} di = \int_0^1 \theta P(j)^{1-\varepsilon} \omega_{t-1}(j) dj = \theta P_{t-1}^{1-\varepsilon}. \tag{B19}$$

Substituting (B19) into (B18) and using the fact that all reoptimizers set $\widetilde{P}_t^{1-\varepsilon}$, we get

$$P_t = \left[(1-\theta)\widetilde{P}_t^{1-\varepsilon} + \theta P_{t-1}^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}. \tag{B20}$$

We divide both sides of (B20) by P_t,

$$1 = \left[(1-\theta)\widetilde{\widetilde{p}}_t^{1-\varepsilon} + \theta\left(\frac{1}{\pi_t}\right)^{1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}},$$

and express $\widetilde{\widetilde{p}}_t$

$$\widetilde{\widetilde{p}}_t = \left[\frac{1-\theta\pi_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}. \tag{B21}$$

Combining (B21) and (B15), we obtain

$$\frac{S_t}{F_t} = \left[\frac{1-\theta\pi_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}. \tag{B22}$$

Aggregate output. Let us define aggregate output

$$\overline{Y}_t \equiv \int_0^1 Y_t(i) di = \int_0^1 \exp\left(\eta_{a,t}\right) L_t(i) di = \exp\left(\eta_{a,t}\right) L_t, \tag{B23}$$

where $L_t = \int_0^1 L_t(i) di$ follows by the labor-market clearing condition. We substitute demand for $Y_t(i)$ from (B5) into (B23) to get

$$\overline{Y}_t = \int_0^1 Y_t \left(\frac{P_t(i)}{P_t}\right)^{-\varepsilon} di = Y_t P_t^{\varepsilon} \int_0^1 P_t(i)^{-\varepsilon} di. \tag{B24}$$

Let us introduce a new variable $\overline{P}_t$,

$$\left(\overline{P}_t\right)^{-\varepsilon} \equiv \int_0^1 P_t(i)^{-\varepsilon} di. \tag{B25}$$

Substituting (B23) and (B25) into (B24) gives us

$$Y_t \equiv \overline{Y}_t \left(\frac{\overline{P}_t}{P_t}\right)^{\varepsilon} = \exp\left(\eta_{a,t}\right) L_t \Delta_t, \tag{B26}$$

where Δ_t is a measure of price dispersion across firms, defined by

$$\Delta_t \equiv \left(\frac{\overline{P}_t}{P_t}\right)^{\varepsilon}. \tag{B27}$$

Note that if $P_t(i) = P_t(i')$ for all i and $i' \in [0,1]$, then $\Delta_t = 1$, that is, there is no price dispersion across firms.

Law of motion for price dispersion Δ_t. By analogy with (B20), the variable $\overline{P}_t$, defined in (B25), satisfies

$$\overline{P}_t = \left[(1-\theta)\,\widetilde{P}_t^{-\varepsilon} + \theta\left(\overline{P}_{t-1}\right)^{-\varepsilon}\right]^{-\frac{1}{\varepsilon}}. \tag{B28}$$

Using (B28) in (B27), we get

$$\Delta_t = \left(\frac{\left[(1-\theta)\,\widetilde{P}_t^{-\varepsilon} + \theta\left(\overline{P}_{t-1}\right)^{-\varepsilon}\right]^{-\frac{1}{\varepsilon}}}{P_t}\right)^{\varepsilon}. \tag{B29}$$

This implies

$$\Delta_t^{\frac{1}{\varepsilon}} = \left[(1-\theta)\left(\frac{\widetilde{P}_t}{P_t}\right)^{-\varepsilon} + \theta\left(\frac{\overline{P}_{t-1}}{P_t}\right)^{-\varepsilon}\right]^{-\frac{1}{\varepsilon}}. \tag{B30}$$

In terms of $\widetilde{\widetilde{p}}_t \equiv \frac{\widetilde{P}_t}{P_t}$, the condition (B30) can be written as

$$\Delta_t = \left[(1-\theta)\,\widetilde{\widetilde{p}}_t^{-\varepsilon} + \theta\frac{\overline{P}_{t-1}^{-\varepsilon}}{P_t^{-\varepsilon}} \cdot \frac{P_{t-1}^{-\varepsilon}}{P_{t-1}^{-\varepsilon}}\right]^{-1}. \tag{B31}$$

By substituting $\widetilde{\widetilde{p}}_t$ from (B21) into (B31), we obtain the law of motion for Δ_t,

$$\Delta_t = \left[(1-\theta)\left[\frac{1-\theta\pi_t^{\varepsilon-1}}{1-\theta}\right]^{-\frac{\varepsilon}{1-\varepsilon}} + \theta\frac{\pi_t^{\varepsilon}}{\Delta_{t-1}}\right]^{-1}. \tag{B32}$$

Aggregate resource constraint. Combining the household's budget constraint (42) with the government budget constraint (53), we have the aggregate resource constraint

$$P_t C_t + P_t \frac{\overline{G} Y_t}{\exp\left(\eta_{G,t}\right)} = (1-v) W_t L_t + \Pi_t. \tag{B33}$$

Note that the ith intermediate-good firm's profit at t is $\Pi_t(i) \equiv P_t(i) Y_t(i) - (1-v) W_t L_t(i)$. Consequently,

$$\Pi_t = \int_0^1 \Pi_t(i)\, di = \int_0^1 P_t(i) Y_t(i)\, di - (1-v) W_t \int_0^1 L_t(i)\, di = P_t Y_t - (1-v) W_t L_t,$$

where $P_t Y_t = \int_0^1 P_t(i) Y_t(i)\, di$ follows by a zero-profit condition of the final-good firms. Hence, (B33) can be rewritten as

$$P_t C_t + P_t \frac{\overline{G}}{\exp\left(\eta_{G,t}\right)} Y_t = P_t Y_t. \tag{B34}$$

In real terms, the aggregate resource constraint (B34) becomes

$$C_t = \left(1 - \frac{\overline{G}}{\exp\left(\eta_{G,t}\right)}\right) Y_t. \tag{B35}$$

Summary. The set of equilibrium conditions derived above can be summarized by six FOCs

$$S_t = \frac{\exp\left(\eta_{u,t} + \eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}} \frac{\left(G_t^{-1} C_t\right)^{1+\vartheta}}{\left(\Delta_t\right)^{\vartheta}} + \beta\theta E_t\left\{\pi_{t+1}^{\varepsilon} S_{t+1}\right\}, \tag{B36}$$

$$F_t = \exp\left(\eta_{u,t}\right) C_t^{1-\gamma} G_t^{-1} + \beta\theta E_t\left\{\pi_{t+1}^{\varepsilon-1} F_{t+1}\right\}, \tag{B37}$$

$$\frac{S_t}{F_t} = \left[\frac{1-\theta\pi_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}, \tag{B38}$$

$$\Delta_t = \left[(1-\theta)\left[\frac{1-\theta\pi_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{\varepsilon}{\varepsilon-1}} + \theta\frac{\pi_t^{\varepsilon}}{\Delta_{t-1}}\right]^{-1}, \tag{B39}$$

$$C_t^{-\gamma} = \beta\frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)} R_t E_t\left[\frac{C_{t+1}^{-\gamma}\exp\left(\eta_{u,t+1}\right)}{\pi_{t+1}}\right], \tag{B40}$$

$$R_t = R_*\left(\frac{R_{t-1}}{R_*}\right)^{\mu}\left[\left(\frac{\pi_t}{\pi_*}\right)^{\phi_\pi}\left(\frac{Y_t}{Y_{N,t}}\right)^{\phi_y}\right]^{1-\mu}\exp\left(\eta_{R,t}\right), \tag{B41}$$

where $Y_{N,t}$ follows from problem (57), (58),

$$Y_{N,t} = \left[\frac{\exp\left(\eta_{a,t}\right)^{1+\vartheta}}{\left[\exp\left(\eta_{G,t}\right)\right]^{-\gamma} \exp\left(\eta_{L,t}\right)} \right]^{\frac{1}{\vartheta+\gamma}}. \tag{B42}$$

Here, S_t and F_t are supplementary variables; Δ_t is a measure of price dispersion across firms. The condition in terms of approximated quantities and approximation errors corresponding to FOCs (B36), (B37), (B38)–(B40) and (B41) are given by (59)–(64), respectively. The problem of minimization of the sum of squared approximation errors is given by (65).

B2. Constructing approximation errors using linearized model's equations

Condition (59). Let us rewrite condition (59) as

$$\frac{\exp\left(\eta_{u,t}+\eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}} \left(G_t^{-1}\widehat{C}_t\right)^{1+\vartheta} (1+\delta_{C_t})^{1+\vartheta} \left(\widehat{\Delta}_t\right)^{-\vartheta} (1+\delta_{\Delta_t})^{-\vartheta}$$
$$+ \beta\theta E_t \left\{ \widehat{\pi}_{t+1}^{\varepsilon} \left(1+\delta_{\pi_{t+1}}\right)^{\varepsilon} \widehat{S}_{t+1} \left(1+\delta_{S_{t+1}}\right) \right\} - \widehat{S}_t \left(1+\delta_{S_t}\right) = 0.$$

Finding a first-order Taylor expansion (again, we use $(1+x)^{\alpha} \simeq 1+\alpha x$) of the last equation and omitting second-order terms, we have

$$\frac{\exp\left(\eta_{u,t}+\eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}} \left(G_t^{-1}\widehat{C}_t\right)^{1+\vartheta} \left(1+(1+\vartheta)\,\delta_{C_t}\right) \left(\widehat{\Delta}_t\right)^{-\vartheta} \left(1-\vartheta\delta_{\Delta_t}\right)$$
$$+ \beta\theta E_t \left\{ \widehat{\pi}_{t+1}^{\varepsilon} \left(1+\varepsilon\delta_{\pi_{t+1}}+\delta_{S_{t+1}}\right) \widehat{S}_{t+1} \right\} - \widehat{S}_t \left(1+\delta_{S_t}\right) = 0.$$

Since $\left(1+(1+\vartheta)\,\delta_{C_t}\right)\left(1-\vartheta\delta_{\Delta_t}\right) = 1-\vartheta\delta_{\Delta_t}+(1+\vartheta)\,\delta_{C_t}$, we have

$$\left[1-\vartheta\delta_{\Delta_t}+(1+\vartheta)\,\delta_{C_t}\right] \frac{\exp\left(\eta_{u,t}+\eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}} \frac{\left(G_t^{-1}\widehat{C}_t\right)^{1+\vartheta} \left(\widehat{\Delta}_t\right)^{-\vartheta}}{\widehat{S}_t}$$
$$+ \left(1+\varepsilon\delta_{\pi_{t+1}}+\delta_{S_{t+1}}\right) \beta\theta E_t \left\{ \widehat{\pi}_{t+1}^{\varepsilon} \frac{\widehat{S}_{t+1}}{\widehat{S}_t} \right\} - 1 - \delta_{S_t} = 0.$$

For convenience, we introduce the following compact notation:

$$\begin{aligned} q_t^1 &\equiv \frac{\exp\left(\eta_{u,t}+\eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}}\frac{\left(G_t^{-1}\widehat{C}_t\right)^{1+\vartheta}\left(\widehat{\Delta}_t\right)^{-\vartheta}}{\widehat{S}_t}, \\ q_t^2 &\equiv \beta\theta E_t\left\{\widehat{\pi}_{t+1}^{\varepsilon}\frac{\widehat{S}_{t+1}}{\widehat{S}_t}\right\}. \end{aligned}$$

Then, we have

$$\left[(\vartheta+1)\,\delta_{C_t}-\vartheta\delta_{\Delta_t}\right]q_t^1+\left(\varepsilon\delta_{\pi_{t+1}}+\delta_{S_{t+1}}\right)q_t^2-\delta_{S_t}+\mathcal{R}_t^1=0, \tag{B43}$$

where $\mathcal{R}_t^1$ is the residual of this FOC, given by (66) in the main text. Introducing compact notation for (B43), we get

$$a_t^{11}\delta_{C_t}+a_t^{13}\delta_{\pi_{t+1}}+a_t^{15}\delta_{\Delta_t}+a_t^{16}\delta_{S_t}+a_t^{17}\delta_{S_{t+1}}+b_t^1=0, \tag{B44}$$

where $a_t^{11}\equiv(\vartheta+1)\,q_t^1$, $a_t^{13}\equiv\varepsilon q_t^2$, $a_t^{15}\equiv-\vartheta q_t^1$, $a_t^{16}\equiv-1$, $a_t^{17}\equiv q_t^2$ and $b_t^1\equiv\mathcal{R}_t^1$.

Condition (60). By finding a first-order Taylor expansion in errors of condition (60) and by taking into account that $\delta_{F_{t+1}}\delta_{\pi_{t+1}}\approx 0$, we obtain

$$\begin{aligned} &\exp\left(\eta_{u,t}\right)G_t^{-1}\widehat{C}_t^{1-\gamma}\left[1+(1-\gamma)\,\delta_{C_t}\right] \\ &\qquad+\beta\theta E_t\left\{\widehat{\pi}_{t+1}^{\varepsilon-1}\widehat{F}_{t+1}\left[1+(\varepsilon-1)\,\delta_{\pi_{t+1}}+\delta_{F_{t+1}}\right]\right\}-\widehat{F}_t-\delta_{F_t}\widehat{F}_t=0. \end{aligned}$$

Introducing compact notations, we get

$$a_t^{21}\delta_{C_t}+a_t^{22}\delta_{F_t}+a_t^{24}\delta_{\pi_{t+1}}+a_t^{28}\delta_{F_{t+1}}+b_2=0, \tag{B45}$$

where $a_t^{21}\equiv(1-\gamma)$ with $q_t^4\equiv\exp\left(\eta_{u,t}\right)G_t^{-1}\frac{\widehat{C}_t^{1-\gamma}}{\widehat{F}_t}$, $a_t^{22}\equiv-1$, $a_t^{24}\equiv(\varepsilon-1)\,q_t^3$ with $q_t^3\equiv\beta\theta E_t\left\{\frac{\widehat{\pi}_{t+1}^{\varepsilon-1}\widehat{F}_{t+1}}{\widehat{F}_t}\right\}$, $a_t^{28}\equiv q_t^3$, $b_t^2\equiv\mathcal{R}_t^2$ and where $\mathcal{R}_t^2$ is the residual in equation (67) in the main text.

Condition (61). A first-order Taylor expansion of (61) yields

$$\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}\left[1+(\varepsilon-1)\,\delta_{\pi_t}\right]}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}\widehat{F}_t\left(1+\delta_{F_t}\right)-\widehat{S}_t\left(1+\delta_{S_t}\right)=0.$$

Equivalently, we write

$$\frac{\widehat{S}_t}{\widehat{F}_t}=\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta}-\frac{\theta\widehat{\pi}_t^{\varepsilon-1}\left[(\varepsilon-1)\,\delta_{\pi_t}\right]}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}\frac{(1+\delta_{F_t})}{(1+\delta_{S_t})}. \tag{B46}$$

The residual of this equation is given in (68). Using the last expression, we write

$$\frac{\widehat{S}_t}{\widehat{F}_t} = \left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}} \left(\mathcal{R}_t^3+1\right)^{-1}. \tag{B47}$$

Using (B47) in (B46), we obtain

$$\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}} \left(\mathcal{R}_t^3+1\right)^{-1} = \left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta} - \frac{\theta\widehat{\pi}_t^{\varepsilon-1}\left[\left(\varepsilon-1\right)\delta_{\pi_t}\right]}{1-\theta}\right]^{\frac{1}{1-\varepsilon}} \frac{\left(1+\delta_{F_t}\right)}{\left(1+\delta_{S_t}\right)}.$$

The latter condition can be rewritten as

$$\left(\mathcal{R}_t^3+1\right)^{-1}\left(1+\delta_{S_t}\right) = 1+\delta_{F_t} - \frac{1}{1-\varepsilon}\frac{\theta\widehat{\pi}_t^{\varepsilon-1}\left(\varepsilon-1\right)\delta_{\pi_t}}{1-\theta\widehat{\pi}_t^{\varepsilon-1}},$$

which yields

$$\underbrace{-}_{\equiv a_{32}}\delta_{F_t} + \underbrace{\frac{\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta\widehat{\pi}_t^{\varepsilon-1}}}_{\equiv a_{33}}\delta_{\pi_t} + \underbrace{\left(\mathcal{R}_t^3+1\right)^{-1}}_{\equiv a_{36}}\delta_{S_t} + \underbrace{\left(\mathcal{R}_t^3+1\right)^{-1}-1}_{\equiv b_3} = 0.$$

Finally, we obtain

$$a_{32}\delta_{F_t} + a_{33}\delta_{\pi_t} + a_{36}\delta_{S_t} + b_3 = 0. \tag{B48}$$

Condition (62). A first-order Taylor expansion of (62) yields

$$\begin{aligned}\left(1-\theta\right)\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{\varepsilon}{1-\varepsilon}}\left[1-\frac{\varepsilon}{\left(\varepsilon-1\right)}\frac{\theta\widehat{\pi}_t^{\varepsilon-1}\left(\varepsilon-1\right)\delta_{\pi_t}}{1-\theta\widehat{\pi}_t^{\varepsilon-1}}\right]\\ +\theta\widehat{\pi}_t^{\varepsilon}\Delta_{t-1}^{-1}\left[1+\varepsilon\delta_{\pi_t}-\delta_{\Delta_t}\right] - \widehat{\Delta}_t^{-1}\left(1-\delta_{\Delta_t}\right) = 0.\end{aligned}$$

In more compact notation, this equation can be written as

$$q_t^5\left[1-\frac{\varepsilon\theta\widehat{\pi}_t^{\varepsilon-1}\delta_{\pi_t}}{1-\theta\widehat{\pi}_t^{\varepsilon-1}}\right] + q_t^6\left[1+\varepsilon\delta_{\pi_t}-\delta_{\Delta_t}\right] - \widehat{\Delta}_t^{-1}\left(1-\delta_{\Delta_t}\right) = 0,$$

where

$$q_t^5 \equiv \left(1-\theta\right)\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta}\right]^{\frac{\varepsilon}{1-\varepsilon}}, \quad q_t^6 \equiv \theta\widehat{\pi}_t^{\varepsilon}\Delta_{t-1}^{-1}.$$

By introducing further simplifying notations, we write

$$a_t^{43}\delta_{\pi_t} + a_t^{45}\delta_{\Delta_t} + b_t^4 = 0, \tag{B49}$$

where

$$a_t^{43} \equiv \varepsilon q_t^6 - q_t^5\frac{\varepsilon\theta\widehat{\pi}_t^{\varepsilon-1}}{1-\theta\widehat{\pi}_t^{\varepsilon-1}}, \quad a_t^{45} \equiv \widehat{\Delta}_t^{-1} - q_t^6, \quad b_t^4 \equiv q_t^5 + q_t^6 - \widehat{\Delta}_t^{-1}.$$

Condition (63). A first-order Taylor expansion of (63) implies

$$\beta\frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)}\widehat{R}_t\left(1+\delta_{R_t}\right)E_t\left\{\exp\left(\eta_{u,t+1}\right)\widehat{C}_{t+1}^{-\gamma}\widehat{\pi}_{t+1}^{-1}\left[1-\gamma\delta_{C_{t+1}}-\delta_{\pi_{t+1}}\right]\right\} - \widehat{C}_t^{-\gamma}\left(1-\gamma\delta_{C_t}\right)=0.$$

This expression can be re-written as

$$\beta\frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)}\widehat{R}_t\left(1+\delta_{R_t}\right)\left[1-\gamma\delta_{C_{t+1}}-\delta_{\pi_{t+1}}\right]E_t\left\{\exp\left(\eta_{u,t+1}\right)\widehat{C}_{t+1}^{-\gamma}\widehat{\pi}_{t+1}^{-1}\right\} - \widehat{C}_t^{-\gamma}\left(1-\gamma\delta_{C_t}\right)=0,$$

which after rearranging the terms becomes

$$\left[1+\delta_{R_t}-\gamma\delta_{C_{t+1}}-\delta_{\pi_{t+1}}\right]\beta\frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)}\frac{\widehat{R}_t}{\widehat{C}_t^{-\gamma}}E_t\left\{\exp\left(\eta_{u,t+1}\right)\widehat{C}_{t+1}^{-\gamma}\widehat{\pi}_{t+1}^{-1}\right\} - \left(1-\gamma\delta_{C_t}\right)=0.$$

The latter equation, combined with the definition of $\mathcal{R}_t^5$ in (70) implies

$$\left[\delta_{R_t}-\gamma\delta_{C_{t+1}}-\delta_{\pi_{t+1}}\right]q_t^7+\gamma\delta_{C_t}+\mathcal{R}_t^5=0,$$

where

$$q_t^7\equiv\beta\frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)}\frac{\widehat{R}_t}{\widehat{C}_t^{-\gamma}}E_t\left\{\exp\left(\eta_{u,t+1}\right)\widehat{C}_{t+1}^{-\gamma}\widehat{\pi}_{t+1}^{-1}\right\}.$$

Introducing further shorter notation, we have

$$a_t^{51}\delta_{C_t}+a_t^{54}\delta_{\pi_{t+1}}+a_t^{59}\delta_{R_t}+a_t^{5,10}\delta_{C_{t+1}}+b_t^5=0, \tag{B50}$$

where

$$a_t^{51}\equiv\gamma,\quad a_t^{54}\equiv-q_7,\quad a_t^{59}\equiv q_7,\quad a_t^{5,10}\equiv-q_7\gamma,\quad b_t^5\equiv\mathcal{R}_t^5.$$

Condition (64). A first-order Taylor expansion of (64) leads us to

$$R_*^{1-\mu}\left(\pi_*^{-\phi_\pi}Y_{N,t}^{-\phi_y}\right)^{1-\mu}\left(R_{t-1}\right)^{\mu}\left[\widehat{\pi}_t^{\phi_\pi}\widehat{C}_t^{\phi_y}G_t^{-\phi_y}\right]^{1-\mu} \times\left[1+\phi_\pi\left(1-\mu\right)\delta_{\pi_t}+\phi_y\left(1-\mu\right)\delta_{C_t}\right]\exp\left(\eta_{R,t}\right) - \widehat{R}_t\left(1+\delta_{R_t}\right)=0.$$

In terms of $\mathcal{R}_t^6$ introduced in (71), the considered equation becomes

$$\left(\mathcal{R}_t^6+1\right)\left[1+\phi_\pi\left(1-\mu\right)\delta_{\pi_t}+\phi_y\left(1-\mu\right)\delta_{C_t}\right]-1-\delta_{R_t}=0.$$

Introducing short notation, we get

$$a_t^{61}\delta_{C_t} + a_t^{63}\delta_{\pi_t} + a_t^{69}\delta_{R_t} + b_t^6 = 0, \tag{B51}$$

where

$$a_t^{61} \equiv \phi_y\left(1-\mu\right)\left(\mathcal{R}_t^6+1\right), \quad a_t^{63} \equiv \phi_\pi\left(1-\mu\right)\left(\mathcal{R}_t^6+1\right), \quad a_t^{69} \equiv -1, \quad b_t^6 \equiv \mathcal{R}_t^6.$$

Minimization problem. The minimization problem with linearized constraints is given by:

$$\begin{aligned}&\min_{x_t}\ \delta_{C_t}^2+\delta_{F_t}^2+\delta_{\pi_t}^2+\delta_{\pi_{t+1}}^2+\delta_{\Delta_t}^2+\delta_{S_t}^2+\delta_{S_{t+1}}^2+\delta_{F_{t+1}}^2+\delta_{R_t}^2+\delta_{C_{t+1}}^2\\&\text{s.t. (B44), (B45), (B48)–(B51).}\end{aligned}$$

Again, we solve this problem to obtain initial guesses on δ's to be used as an input for a solver that finds approximation errors from nonlinear FOCs.

B3. Constructing approximation errors using nonlinear model's equations

Instead of linearizing FOCs, we now focus on their nonlinear (original) versions. To simplify the analysis, we obtain (wherever possible) explicit conditions for approximation errors. In particular from six equations (59)–(64), we get explicit formulas for $\delta_{\Delta_t}^2$, $\delta_{S_t}^2$, $\delta_{S_{t+1}}^2$, $\delta_{F_{t+1}}^2$, $\delta_{R_t}^2$, $\delta_{C_{t+1}}^2$. Specifically, we have
– from (62)

$$\delta_{\Delta_t} = \frac{1}{\widehat{\Delta}_t}\left[(1-\theta)\left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}\left(1+\delta_{\pi_t}\right)^{\varepsilon-1}}{1-\theta}\right]^{\frac{\varepsilon}{\varepsilon-1}} + \theta\frac{\widehat{\pi}_t^{\varepsilon}\left(1+\delta_{\pi_t}\right)^{\varepsilon}}{\Delta_{t-1}}\right]^{-1} - 1; \tag{B52}$$

– from (61)

$$\delta_{S_t} = \left[\frac{1-\theta\widehat{\pi}_t^{\varepsilon-1}\left(1+\delta_{\pi_t}\right)^{\varepsilon-1}}{1-\theta}\right]^{\frac{1}{1-\varepsilon}}\frac{\widehat{F}_t}{\widehat{S}_t}\left(1+\delta_{F_t}\right) - 1; \tag{B53}$$

– from (59)

$$\delta_{S_{t+1}} = \frac{\widehat{S}_t\left(1+\delta_{S_t}\right) - \frac{\exp\left(\eta_{u,t}+\eta_{L,t}\right)}{\left[\exp\left(\eta_{a,t}\right)\right]^{\vartheta+1}}\frac{\left(G_t^{-1}\widehat{C}_t\right)^{1+\vartheta}\left(1+\delta_{C_t}\right)^{1+\vartheta}}{\left(\widehat{\Delta}_t\right)^{\vartheta}\left(1+\delta_{\Delta_t}\right)^{\vartheta}}}{\left(1+\delta_{\pi_{t+1}}\right)^{\varepsilon}\beta\theta E_t\left\{\widehat{\pi}_{t+1}^{\varepsilon}\widehat{S}_{t+1}\right\}} - 1; \tag{B54}$$

– from (60)

$$\delta_{F_{t+1}} = \frac{\widehat{F}_t\left(1+\delta_{F_t}\right) - \exp\left(\eta_{u,t}\right) G_t^{-1}\widehat{C}_t^{1-\gamma}\left(1+\delta_{C_t}\right)^{1-\gamma}}{\left(1+\delta_{\pi_{t+1}}\right)^{\varepsilon-1}\beta\theta E_t\left\{\widehat{\pi}_{t+1}^{\varepsilon-1}\widehat{F}_{t+1}\right\}} - 1; \tag{B55}$$

– from (64)

$$\begin{aligned}\delta_{R_t} &= R_*^{1-\mu}\left(\pi_*^{-\phi_\pi}Y_{N,t}^{-\phi_y}\right)^{1-\mu}\left(\frac{R_{t-1}}{\widehat{R}_t}\right)^{\mu}\left[\widehat{\pi}_t^{\phi_\pi}\left(G_t^{-1}\widehat{C}_t\right)^{\phi_y}\frac{1}{\widehat{R}_t}\right]^{1-\mu} \\ &\quad\times\exp\left(\eta_{R,t}\right)\left[\left(1+\delta_{\pi_t}\right)^{\phi_\pi}\left(1+\delta_{C_t}\right)^{\phi_y}\right]^{1-\mu} - 1;\end{aligned} \tag{B56}$$

– from (63)

$$\delta_{C_{t+1}} = \left[\frac{\left(1+\delta_{C_t}\right)^{-\gamma}\left(1+\delta_{\pi_{t+1}}\right)\widehat{C}_t^{-\gamma}}{\left(1+\delta_{R_t}\right)\frac{\exp\left(\eta_{B,t}\right)}{\exp\left(\eta_{u,t}\right)}\widehat{R}_t\beta E_t\left[\frac{\widehat{C}_{t+1}^{-\gamma}\exp\left(\eta_{u,t+1}\right)}{\widehat{\pi}_{t+1}}\right]}\right]^{-\frac{1}{\gamma}} - 1. \tag{B57}$$

We substitute the above expressions into the objective function (65). As a result, we have a minimization problem with four unknowns, $\delta_{C_t}^2$, $\delta_{F_t}^2$, $\delta_{\pi_t}^2$, $\delta_{\pi_{t+1}}^2$

$$\min_{\delta_{C_t}^2,\delta_{F_t}^2,\delta_{\pi_t}^2,\delta_{\pi_{t+1}}^2} \delta_{C_t}^2+\delta_{F_t}^2+\delta_{\pi_t}^2+\delta_{\pi_{t+1}}^2+\underbrace{\delta_{\Delta_t}^2+\delta_{S_t}^2+\delta_{S_{t+1}}^2+\delta_{F_{t+1}}^2+\delta_{R_t}^2+\delta_{C_{t+1}}^2}_{\text{given by (B52)–(B57)}}. \tag{B58}$$

This problem was solved to produce the results in Table 5.