Dynamic Perturbation*

Alessandro Mennuni†  Serhiy Stepanchuk‡
May 4, 2022

Abstract: We develop a new algorithm to solve large-scale macroeconomic models over a large transition. The method consists of Taylor expanding the equilibrium conditions of the model not just around a steady state, but sequentially along the entire equilibrium path. The method can be applied to a broad class of models and is orders of magnitude more accurate than solutions based on local perturbation of the steady state. The method is also able to solve models with strong nonlinearities. Finally, because our policies are locally linear, we can make use of a version of the Kalman filter with time varying coefficients to identify shocks from data. With this tool in hand we are able to evaluate the likelihood function which can be used for the estimation of nonlinear models.

Keywords: Large-Scale Models, Nonlinearities, Transition, Business Cycle, Computational Economics

JEL: E32, C63

*We thank Martin Gervais, Michael Hatcher, Michal Kejak, Paul Klein, Felix Kubler, Yang Lu, and seminar participants at CEF 2018, Global Macro Workshop CERGE-EI 2019, University of Manchester, University of Zurich, University of Miami, University of Lanchaster, and University of Portsmouth.
†University of Southampton. alessandro.mennuni@soton.ac.uk
‡University of Southampton. s.stepanchuk@soton.ac.uk
1 Introduction

This paper develops an algorithm to compute and estimate large-scale and highly non-linear dynamic macroeconomic models, including the transitional dynamics after policy, demographic, technological changes, or big shocks.

When solving dynamic macroeconomic models, a key distinction is between global methods that solve the model over the entire state space, and local methods that solve the model at a steady state.\textsuperscript{1} The method that we develop lies in between. In particular, we develop a way to Taylor expand the policy functions of a model away from a steady state. With this innovation in hand it is then possible to apply local approximations repeatedly along an equilibrium path. As a result, our solution is more accurate than the one found with standard perturbation methods. At the same time, compared with global methods, ours is faster and can solve larger models, and is comparable in terms of accuracy along the simulated path. Furthermore, we can use the solution produced by our algorithm to estimate nonlinear models based on maximum likelihood and Bayesian methods: since with our method the policy functions are approximated by locally linear functions, we can use an extension of the usual Kalman filtering techniques with time varying coefficients that capture models’ nonlinearities and nonstationarity.

We provide an intuitive and a formal description of the algorithm in section 3. In addition to being able to apply Taylor expansion away from a steady state, a further contribution relative to the standard perturbation methods is that our method allows us to explicitly incorporate expectations and deal with the Jensen’s inequality. Perturbation methods assume certainty equivalence at the first order but capture the impact of future uncertainty through the higher order terms approximated around the deterministic steady state. Instead we explicitly model the future probability distribution by incorporating

next period’s states of nature. This is possible because our method iterates backwards from the steady state and gives us future policy functions which can be evaluated at different possible future states of nature. This way to approximate expectations enables us to approximate richer stochastic environments and resembles the one taken in global methods such as value or policy function iteration.

To highlight the strengths of our algorithm, we apply it in 3 different models that pose challenges to existing numerical solutions: a multicountry neoclassical model with capital irreversibility, a New Keynesian model with rare disaster shocks and a zero lower bound constraint on the nominal interest rate, and a model with occasionally binding borrowing constraints.

Our first application, a multicountry neoclassical growth model, is a standard setup that is often used to test numerical algorithms that can handle high-dimensional models, because it is easily scalable by choosing the number of countries.\(^2\) Furthermore, following Brumm & Scheidegger (2017), we add a capital irreversibility constraint that is occasionally binding. This feature generates a strong nonlinearity in the policy functions. This highlights the ability of our algorithm to be accurate even in the presence of such nonlinearities. To get a sense of scalability, we solved the model with 100 countries (200 state variables) with the irreversibility constraint, and 200 without. The time increases with the number of countries but far less than exponentially. In comparison, Brumm & Scheidegger (2017) solve the model for 10 countries. They get to 50 in the smooth model without the irreversibility constraint. However, a limitation of our approach is that we do not compute global policy functions. Rather, we solve for a given equilibrium path and our solution time is roughly proportional to the number of periods in the simulation.

As mentioned, a further contribution of our method relative to the standard perturbation approaches is that it can explicitly incorporate expectations and deal with the Jensen’s inequality. This motivates our second application — a New Keynesian model

\(^2\)See Kollman et al. (2011).
with recursive preferences, capital and rare disasters. As argued by Fernandez-Villaverde & Levintal (2018), models with rare disaster allow one to address asset pricing puzzles because of the “large precautionary behaviour responses that are induced in normal times by the probability of the tail events”. They also argue that this makes it an especially challenging model to compute numerically. Similarly to the methods used by Fernandez-Villaverde & Levintal (2018), our approach delivers a high equity premium and a low risk free rate, in line with empirical evidence. Because the results rely on tail events, they cannot be generated by higher order perturbation of the steady state.

Inspecting the nominal implications of the New Keynesian model above, we noticed that with the level of tail risk needed to match the evidence on asset prices, the nominal interest rate is negative. This led us to add to the model the zero-lower bound (ZLB). This feature adds non-linearities and has recently received attention from the computational literature, but we are not aware of any paper that has combined capital, rare disasters and the ZLB. We show that the addition of the ZLB increases the risk-free rate and reduces the equity premium. This finding highlights some challenges that models with nominal frictions pose for the disaster shock mechanism to address asset pricing puzzles.

Finally, we show how our method can be combined with an extension of the Kalman filter which allows us to estimate nonlinear models. In particular, since the policies provided by our solution method are locally linear, we can use these local linear approximations to the policy and transition functions as inputs for the extended Kalman filter (EKF), the nonlinear version of the KF that sequentially linearizes the equations of the model around the best current estimate of the model’s state (see Anderson & Moore 1979). As a laboratory model for this section we use Mendoza (2010) who develops a model with occasionally binding financial constraints that generate “sudden stops” in emerging economies. This is an especially useful application to highlight the merits of our approach because the borrowing constraint does not bind in steady state but is occasionally binding over the equilibrium path causing sudden stops. As a result, with the usual KF, a
key parameter that regulates the tightness of the financial constraint is not identified: because in the steady state the financial constraint does not bind, the policy functions approximated there are largely independent of the tightness of the financial constraint. And since the linear KF uses policy functions linearized around that point, the resulting likelihood function is flat with respect to the key parameter that determines borrowing capacity. Instead we show that our method identifies this parameter correctly.

Our paper proceeds as follows: section 2 summarizes the literature, section 3 describes the algorithm, section 4 applies it to the multi-country model with capital irreversibility, section 5 applies it to a model with rare disaster shocks and zero lower bound, section 6 illustrates the KF application for model estimation, section 7 concludes.

## 2 Related literature

Our paper is related to several strands of literature in economics and in other computational fields. In particular, outside economics there have been exciting developments in the field of optimal control which are related to the algorithm we propose. In a recent survey, Bertsekas (2021) argues that the most important recent innovation is the introduction of what he calls online, reinforced learning algorithms where most of the computation is performed sequentially: as with our approach, the solution is computed only at the relevant states as they become known (hence the name “online”) and not all over the state space as with global ("offline") methods. A prominent example of an online scheme is Model Predictive Control (MPC). However, this approach is designed to solve optimal control problems and it is not applicable to problems whose solution is not Pareto efficient and cannot be expressed as a planning problem. Instead our approach deals with a general equilibrium system of difference equations, including those that cannot be cast as a solution of a planning problem. Furthermore, the MPC algorithm uses certainty

---

3This approach solves at each $t$ a finite truncated version of the model and retains only the solution for the first period. Then, after observing the shock, the process is repeated. See Grüne et al. (2015) for an economic application of MPC.
equivalence or a highly simplified future stochastic process to avoid the proliferation of future nodes, see Rawlings et al. (2019) chapter 3.4

Our approach is also related to the Expected Path algorithm developed by Fair & Taylor (1983) where, like us, they solve for a given equilibrium path. Two key advantages of our approach are that it is more scalable and we can handle stochastic environments without assuming certainty equivalence. Our scalability advantage comes from the fact that Fair & Taylor (1983), similarly to MPC, solve the equilibrium conditions simultaneously for the whole finite truncated future path. We break this into a sequence of two-period problems making use of the next period policy functions.5

Anderson et al. (2012) study small noise expansions for control problems with recursive risk-sensitive preferences. Like us, their algorithm exploits a deterministic path from the current value of the state vector towards the steady state. However, our algorithms differ both in the way we construct and utilize such path. In particular, they compute an expansion of the value function with respect to a parameter that scales the Gaussian shocks; this is similar to the perturbation approach from the control-theory literature, which perturbs the solution in the parameter space as opposed to the state space, see Simmonds (1998). Instead, we Taylor expand around the equilibrium points along the path and use these expansions to update our approximation to the next-period policy functions, which we then use to solve the non-linear equilibrium system. As a result, we do not perturb over small Gaussian noise but, similarly to policy function iteration, we can consider richer stochastic environments such as rare disaster shocks and other strong nonlinearities such as occasionally binding constraints.6 Finally, they work with control

---

4See Bertsekas (2021) for other related online algorithms, such as those that contributed to the big breakthrough in the modelling of 2 player, discrete choice games including the celebrated Alpha-Zero algorithm for playing chess. Even though the applicability is different from ours, they are prominent examples of how online steps can greatly improve performance over offline algorithms.

5There are a few papers that build on the Extended path of Fair & Taylor (1983) and go beyond the assumption of certainty equivalence such as Adjemian & Juillard (2013) and Maliar et al. (2020). However, they do not address the aforementioned curse of dimensionality.

6On the other hand, through their expansion approach, Anderson et al. (2012) obtain analytic formulas that characterize the impact of uncertainty and risk sensitivity on economic decisions and asset prices.
problems while we work with the equilibrium conditions. As a result, we can also solve models with inefficient equilibria.

Closer to the standard perturbation approach used in economics, Den Haan et al. (2016) and Evans & Phillips (2017) try to extend perturbation methods around an equilibrium point away from the steady state. Like us, they Taylor expand the system of equilibrium conditions of the model. But their approach differs from ours in a crucial aspect. In principle the Taylor expansion can be done at any point that satisfies the equilibrium conditions, not just at the steady state. However, there is an issue: there are many solutions to the system of equilibrium equations, not all of which are also solutions of the underlying economic problem. To find policy functions from the system of equilibrium conditions requires further restrictions. The approach in a steady state is to restrict the policy functions to rule out explosive solutions (see Blanchard & Kahn 1980). The same approach is used also away from the steady state by Evans & Phillips (2017). However, as they acknowledge, true policies may not be locally stable outside of the steady state so this method can be erroneous. As an alternative restriction, Den Haan et al. (2016) assume that the policy functions for the current and the next period are the same.\footnote{A similar restriction is adopted in Levintal (2018), who combines the Taylor expansion of the equilibrium system with projection methods.} Again this assumption may not be warranted when away from a steady state, especially if policies are highly nonlinear. Our approach keeps current and future policies distinct. As further restrictions needed to identify the solution we use the next period policy functions, which we obtain by iterating backwards from the steady state.

3 Description of the algorithm

In this section we describe our numerical algorithm. We introduce some notation in section 3.1. We then provide an intuitive outline of the algorithm in section 3.2. We provide a more detailed step-by-step description in section 3.3. Sections 3.4 and 3.5
contain theoretical results about the accuracy of our method.

3.1 The general setup

Following Schmitt-Grohe & Uribe (2004) and Gomme & Klein (2011) among others, we consider a dynamic general equilibrium model that can be formulated as a system of equilibrium conditions:

$$E_t[f(x_{t+1}, y_{t+1}, x_t, y_t)] = 0,$$

and appropriate transversality conditions. In Equation (1) $E_t$ is the expectation given information at time $t$, $x_t$ is a vector of size $n_x$ of current period realizations of the predetermined (or state) variables, $y_t$ is a vector of size $n_y$ of current period realizations of the non-predetermined (or “jump”) variables of the model, while $x_{t+1}$ and $y_{t+1}$ are their corresponding next-period realizations.

The state vector $x_t$ can be partitioned into:

$$x_t = [x^1_t, x^2_t],$$

where $x^1_t$ consists of endogenous state variables, while $x^2_t$ follows some exogenous stochastic process. $x^2_t$ follows a stationary VAR(1) process:

$$x^2_{t+1} = \Lambda x^2_t + \eta \varepsilon_{t+1}$$

where $\varepsilon_t$ is a vector of shocks (of size $n_{\varepsilon} = n_{x^2}$) that have zero mean and variance matrix $I$, and $\eta$ is an $n_{\varepsilon} \times n_{\varepsilon}$ matrix. We use this linear representation of the stochastic process in all our applications, but our algorithm can work with non-linear versions of (3).

Our aim is to obtain an equilibrium path given an initial value of the state variables, $x_0$, and a finite sequence of shocks, $\{\varepsilon_t\}$. Notice that the sequence of shocks is known to the modeller; the agents in the model form expectations over the future shocks.

The vector-valued function $f$ typically combines first-order static and dynamic op-
timality conditions that characterize optimal choices of economic agents populating the model, market-clearing conditions and the equations that characterize the laws of motion for the endogenous and exogenous state variables. It consists of \( n = n_x + n_y \) possibly non-linear equations.

A recursive representation of the solution takes the form of the policy function for the jump variables:

\[
y_t = g(x_t)
\]  

(4)

and the transition function for the state variables:

\[
x_{t+1} = h(x_t) + \eta \varepsilon_{t+1}
\]  

(5)

where \( \eta = [0, \tilde{\eta}]' \).

### 3.2 An intuitive description of the algorithm

The key insight in our method is that one can construct a Taylor series approximation to the system of equilibrium conditions at any dynamic point \((x_{t+1}, y_{t+1}, x_t, y_t)\) that satisfies this system, not just the steady state. Here we offer an intuitive and non-technical discussion. For simplicity, we start with the deterministic version of our algorithm.

The first thing to notice is that not every point \((x_{t+1}, y_{t+1}, x_t, y_t)\) that solves the equilibrium system is on the equilibrium path. For instance, there may be points that solve the equilibrium system but are on an explosive path and would violate transversality conditions. Taylor expanding the equilibrium system of equations at such unstable points may lead to policies that poorly approximate the equilibrium ones. One way to see this difficulty is to note that \( f \) in the equilibrium system (1) provides only \( n = n_x + n_y \) equations, while we need to find the solution for \( n_x + n_y + n_y \) values of \((x_{t+1}, y_t, y_{t+1})\) (the \( n_x \) values of \( x_t \) are predetermined and known in period \( t \)). Since there are more

---

8We borrow this notation from Schmitt-Grohe & Uribe (2004). The first element in \( \eta \) is a zero matrix associated with the endogenous state variables, \( x_1^t \) in (2), which are not directly affected by the exogenous innovations. The exogenous innovations are impulses to the exogenous state variables \( x_2^t \) via matrix \( \tilde{\eta} \).
unknowns than equations, \( f \) does not allow us to find the unique equilibrium values for 
\((x_{t+1}, y_t, y_{t+1})\) given \( x_t \). Therefore, a key step in our algorithm is to be able to start with 
any initial value of the state vector \( x_0 \) (potentially far from the steady state), and find 
the corresponding values of \( y_0, x_1 \) and \( y_1 \).

The problem of finding \((x_1, y_1, y_0)\) given \( x_0 \) would be solved if we had the policy 
function \( g \) defined in (4) to determine \( y_1 \). This would solve the problem because it would 
give us \( n_y \) further conditions \( y_1 = g(x_1) \), which, together with the \( n_x + n_y \) equations 
given by (1), pin down a unique \((x_1, y_1, y_0)\). Intuitively, the policy \( g \) gives us the optimal 
future path starting from any \( x_1 \) and thus puts us on the equilibrium path. Our algorithm 
provides an efficient way to get an accurate local solution to future \( g \), through which we 
can solve for \((x_1, y_1, y_0)\).

To illustrate, consider a simple example of the deterministic neoclassical growth model 
where the equilibrium is described by the Euler equation:

\[
\frac{u'(c_t)}{\beta u'(c_{t+1})} = \alpha A k_{t+1}^{\alpha-1} + 1 - \delta
\]

and the resource constraint:

\[
k_{t+1} + c_t = Ak_t^\alpha + (1 - \delta)k_t.
\]

where the utility function \( u \) has the usual neoclassical properties and \( 0 < \delta \leq 1, 0 < \beta < 1, \) 
\( 0 < \alpha < 1 \) and \( A > 0 \).

In terms of notation, the 2 equations above form \( f(\cdot) \). The state variable \( x_t \) is capital, 
\( k_t \), and the jump variable \( y_t \) is consumption, \( c_t \).

Suppose \( t = 0 \). If we know the policy function for consumption in period \( t = 1, \)

\[\text{A steady state solves this problem because of the further steady state restrictions that } x_t = x_{t+1} \text{ and } y_t = y_{t+1}.\]
\(c_1 = g(k_1)\), we can substitute it into the Euler equation above and get

\[
u'(c_0) = \beta u'(g(k_1)) \left( \alpha A k_1^{a-1} + 1 - \delta \right).
\] (6)

The latter and the resource constraint are 2 equations in the 2 unknowns \(c_0\) and \(k_1\) and can be easily solved given \(k_0\). We do not have \(g\); however, we know that around the deterministic steady state, the linear approximation of \(g\) around the steady state is fairly accurate. Thus it can work well if \(k_1\) is close to the steady state (or if the policy is fairly linear). We call this way to solve for \(c_0\) and \(k_1\) using an approximation to future \(g\) our one-step approach. It is similar to a one-time policy iteration, but applied in a single point of the state space rather than on the whole (or a large part) of it. Later we prove theoretically that for any admissible initial \(k_0\), this approach is better than applying the steady state policies at \(k_0\) to compute \(c_0\) and \(k_1\).

**Multi-Step**

If the initial point \(k_0\) is very close to the steady state, so that the model converges to a small neighbourhood of the steady state in one period, then it makes sense to use the steady state policy for next period consumption. But what if \(k_0\) was not that close to the steady state? Suppose it was slightly further away so that it took not one but 2 periods to reach the neighbourhood of the steady state. In this case the Taylor approximation of \(g\) around the steady state may not be an accurate policy for next period consumption (because \(k_1\) is far from steady state capital \(k_{ss}\), but it would be accurate in 2 periods time, when \(k_2 \approx k_{ss}\). So, we can apply the procedure above twice:

1) at some intermediate point \(\tilde{k}_1\) we apply the one-step procedure using steady state future policy;

2) derive the Taylor approximation of the policy function for consumption around this intermediate point;

3) repeat the one-step procedure at \(k_0\), using the approximation to \(g\) around the interme-
diate point instead of the steady state approximation.

One issue that remains is how to pick the intermediate point $\tilde{k}_1$. In this one-dimensional example it is fairly intuitive that it should lie between $k_0$ and the steady state $k_{ss}$, but it is not so obvious where exactly it should be. There are potentially many approaches to choose $\tilde{k}_1$, but one that is simple and proved to work well in all our applications is to first simulate the model from $k_0$ to $k_{ss}$ with the steady state policy functions assuming no shocks. This gives us an auxiliary path $\{k_0, \tilde{k}_1, k_{ss}\}$ on which we can proceed backwards as in points 1, 2 and 3 above.\(^\text{10}\) Finally, if $k_0$ is further away from the steady state, the auxiliary path can take more than 3 points; as many as necessary to get sufficiently close to the steady state.\(^\text{11}\) An example of a typical auxiliary path is shown in figure 1, left panel.

The simple example above assumed a deterministic model where capital is the only state variable. The approach also works in a stochastic model if one is satisfied with the certainty equivalence assumption. We just need to add to $f(\cdot)$ a third equation, the law of motion of productivity $A$ such as $A_{t+1} = \rho A_t + \eta \varepsilon_t$, and replace $A$ in the right-hand-side of equation (6) with $E_t(A_{t+1}) = \rho A_t$. Now $x_t = [k_t, A_t]$. Alternatively, we can capture the effect of the higher moments of the shocks in a way which is similar to how it is done with policy iteration, as described later.

Notice that the procedure so far finds the solution at $x_0$. We can then use the policies approximated around $x_0$ and the realization of the shocks $\varepsilon_1$ to obtain $y_0$ and $x_1$, the elements of the stochastic equilibrium path we are solving for. To obtain the next point in the stochastic simulation given a time series for the shocks, we repeat the above procedure.

\(^{10}\)Another way to get the auxiliary path is to solve a perfect foresight path as in Fair & Taylor (1983).

\(^{11}\)Boppart et al. (2018) and Auclert et al. (2021) also use an auxiliary perfect foresight path towards the steady state to solve models with a continuum of heterogeneous agents. Our algorithm is not specifically designed to solve models with a continuum of agents like theirs. In addition to the different economic environment, a difference is that they Taylor expand in sequence space (i.e. linearizing around the entire path) whereas we Taylor expand iteratively and backwards, point by point, along this path.
starting with $x_1$ as the initial point.\footnote{I.e. we draw a new auxiliary path from $x_1$ to the steady state, and follow it backward to find the policies at $x_1$. In doing so, we discard all the local policies found on the previous auxiliary path. To speed up the code, one could avoid drawing new auxiliary paths by recycling the local policies on the initial auxiliary path. Those are not perfectly aligned with $x_1$ but may provide a good speed-accuracy compromise.} This can be repeated to produce a simulated equilibrium path of any desired length. Figure 1, right panel, illustrates how the state vector $x_t$ in the first 3 periods is obtained through this procedure.\footnote{To avoid a 3-dimensional graph, the panel is plotting a linear combination of the two state variables $k$ and $A$ on the vertical axis.}

\subsection{The algorithm details}

In this section, we provide a more detailed step-by-step description of our numerical algorithm. The following additional notation will be useful: let $g_{ss}(\cdot)$ denote the Taylor approximation to $g$ around the steady state, and let $g_\hat{x}(\cdot)$ be the one around some generic point $\hat{x}$.\footnote{To be more precise, the point of approximation is $(\hat{x}', \hat{y}', \hat{x}, \hat{y})$ that solves \eqref{eq:1}, where we use primes to denote next period variables. But since our procedure gives us $\hat{y}$, $\hat{x}'$ and $\hat{y}'$ for a given $\hat{x}$, for ease of notation we use only $\hat{x}$ to denote the point of approximation.}

We start with formally defining our goal.

**Objective:**

The aim is to obtain an equilibrium path $\{(x_0, y_0), (x_1, y_1), ..., (x_T, y_T)\}$ given an initial value of the state variables, $x_0$, and a sequence of realized shocks, $\{\varepsilon_t\}_1^T$, where $T$ is the length of the shocks’ time series.
We start with a Taylor approximation for the policy functions around the stable steady state, \( h_{ss}(x) \), \( g_{ss}(x) \), where \((x_{ss}, y_{ss})\) solve \( f(x_{ss}, y_{ss}, x_{ss}, y_{ss}) = 0 \). This preliminary step can be replaced with some alternative initial approximation to the policy functions.\(^{15}\)

If the steady state approach is used, the resulting policies must be locally stable (for stability, see for instance Blanchard & Kahn 1980).\(^{16}\) Our algorithm below refines these initial approximations.\(^{17}\)

We start with describing the algorithm under certainty equivalence.

1. Put \( t = 0 \) and set \( x_t = x_0 \).

2. Set the shocks to 0. Using the steady state policy \( h_{ss} \), generate the auxiliary path \( \{x_\tau\}_{\tau=0}^{T_{aux}} \), starting with \( x_0 = x_t \) and where \( T_{aux} \) is the length of the auxiliary path.\(^{18}\), \(^{19}\)

\( \text{Step (2) corresponds to the construction of the auxiliary path (the dashed line) in the left panel of figure 1. The following steps (3-6) trace the auxiliary path backwards from the steady state to the current state} \ x_t. \text{The purpose of the auxiliary path is to sequentially update the local approximation of} \ g; \text{this is similar to policy iteration but performed on a single path.} \)

3. Set \( \tau = T_{aux} \), and initiate \( g_{\bar{x}}(\cdot) \) with the steady state approximation, \( g_{ss}(\cdot) \).

\(^{15}\)If a sufficiently good initial approximation to the policy functions is available globally, tracing the auxiliary path backwards may be unnecessary and our one-step algorithm can work well.

\(^{16}\)Stability is necessary because it ensures that the auxiliary path converges to the steady state. Stability does not imply uniqueness: if the model has indeterminacy of the steady state policy functions, then one has to be selected.

\(^{17}\)As mentioned, following recent developments in optimal control and reinforced learning, one can distinguish between off-line and on-line computations (see Bertsekas (2021)). Our algorithm below corresponds to the on-line step where the computations are performed sequentially, as the current state becomes known, so the policy functions are computed only at the relevant states. One can view the steady state approximation as the off-line step.

\(^{18}\)Ideally \( T_{aux} \) should be long enough so that \( x_{T_{aux}} \) is close to \( x_{ss} \). But we show theoretically that even \( T_{aux} = 1 \) (our one step algorithm) makes an improvement over steady state linearization and we confirm numerically that it can be quite accurate.

\(^{19}\)Notice that we index the auxiliary path with \( \tau \), which is different from the index of the equilibrium path, \( t \).
(4) Find \( x' \) and \( y \) such that \( f(x', g(x'), \bar{x}_\tau, y) = 0 \) (note that we substituted \( y' \) with \( g(x') \), so we solve the \( n_x + n_y \) equations in \( f(\cdot) = 0 \) for the \( n_x + n_y \) unknowns \( x', y \)).

(5) Apply the implicit function theorem to \( f(x', g(x'), \bar{x}_\tau, y) = 0 \) to derive a new Taylor approximation of the policy functions. Unlike at the steady state, this step is a straightforward application of the implicit function theorem because, owing to the future policy function, we do not need to also solve for \( y' \) and so we have as many equations as unknowns; see Appendix A. Update the policy function: \( g(x') = g(x', \cdot) \).

(6) If \( \tau > t \), set \( \tau = \tau - 1 \) and go back to step (4).

Finally, steps (7) and (8) build the stochastic simulation. This corresponds to the movements along the solid blue line in the right panel of figure 1.

(7) If \( \tau = t \), we have found a local solution around the current value of the state \( x_t \). We then update the equilibrium sequence defined in our objective with \( y_t = g(x_t) \) and \( x_{t+1} = h(x_t) + \eta \varepsilon_{t+1} \).

(8) If \( t = T \), the whole simulated equilibrium sequence has been found! Otherwise, set \( t = t + 1 \) and go back to step (2).

3.3.1 Modelling Expectations

It is possible to modify the algorithm above to avoid assuming certainty equivalence.

Replace step (4) with the following:

4(b) Find \( x', y', \) and \( y \) such that \( \hat{E}_\tau f(x', y', \bar{x}_\tau, y) = 0 \), where \( y' = g(x') \) for all realizations of \( x' \), and where \( \hat{E}_\tau \) is an approximation to the conditional expectations, for instance using a quadrature method.

With the expectation operator outside \( f(\cdot) \), \( x' \) and \( y' \) are no longer unknowns of size \( n_x \) and \( n_y \) respectively; we need to solve for the state contingent values given the
future shocks’ realizations embedded in \( x' \). So there are now more unknowns than in the certainty equivalence case (see section 3.2). As a result, we need more restrictions. The state contingent values of \( x' \) are easily determined by the exogenous law of motion of the shock.\(^{20}\) The hard part is how to determine \( y' \) which is endogenous and state contingent. However, the future policy \( g_{\hat{x}}(\cdot) \) gives us \( y' = g_{\hat{x}}(x') \) for all future possible shock realizations. So, with the future policy \( g_{\hat{x}}(\cdot) \) in hand, we still only need to solve for \( x' \) and \( y \) (i.e. we substitute \( y' \) out and solve \( \hat{E}_T f (x', g_{\hat{x}}(x'), x', y) = 0 \)). This is not in general possible without knowing future \( g \). Let us illustrate this in the simple RBC example that we have used previously in section 3.2. Suppose that we approximate expectations using an \( m \)-point quadrature. Then the equilibrium system of equations consists of the resource constraint and the following Euler equation:

\[
u'(c) = \beta \left( \sum_{j=1}^{m} \omega_j u'(g_k(k', \rho A + \eta \epsilon_j)) \left( \alpha(\rho A + \eta \epsilon_j)(k')^{\alpha-1} + 1 - \delta \right) \right) \tag{7}\]

where \( m \) is the number of quadrature points \( \epsilon_j \), while \( \omega_j \) are the corresponding quadrature weights.\(^{21}\) The quadrature points do not enter the resource constraint, where only the current productivity value appears. As a result, the unknowns are \( c \) and \( k' \) which are not contingent on the future shocks and therefore are not indexed by the quadrature points. Thus this is a system of 2 equations in 2 unknowns. The same approach of using the future state-contingent policy function inside the expected value is used in the policy iteration approach, and it accounts for the Jensen’s inequality: in general, the right hand side of equation 7 is different than its certainty equivalence counterpart

\[
\beta \left( u'(g_k(k', \rho A)) \left( \alpha(\rho A)(k')^{\alpha-1} + 1 - \delta \right) \right).
\]

\(^{20}\) \( x' = [x'^1, x'^2] \) is state contingent only in its second component \( x'^2 \). The latter is easily determined state by state by the law of motion (3).

\(^{21}\) Note that in the Euler equation above we substituted out the state contingent part of \( x' \), i.e. \( A'_j \), with its law of motion.
That we account for the Jensen’s inequality may seem surprising given the well-known finding of Schmitt-Grohe & Uribe (2004), that first order perturbation around the steady state does not account for the effects of the variance of the shocks. The difference is that they

(i) solve the deterministic system for a steady state,
(ii) derive the policy functions around such point.

As a result, the solution to (i), being deterministic, is not affected by the Jensen’s inequality. It follows that the variance of the shocks does not enter the policies derived by first order Taylor expansion of the system as shown by Schmitt-Grohe & Uribe (2004). This result does not apply to us because instead of (i) we solve the stochastic and nonlinear system given future linear policy functions. This gives us solutions $y, x'$ that are affected by the shocks’ higher moments. In turn, this affects the constant term and the slope of the policy functions approximated around that solution.\footnote{The aforementioned result of Schmitt-Grohe & Uribe (2004) that, up to the first order, the policies do not directly depend on the variance of the shocks, still applies. But they depend on it indirectly, as it affects the point of approximation and thereby the constant and slope of the policies.}

Despite the fact that we are solving the state contingent model, the reader may still be worried that our use of linear future policy function could undo the Jensen’s inequality. We illustrate with a simple example that a future linear policy function does not eliminate the Jensen’s inequality. To this aim, consider the standard example of precautionary savings in a 2 period savings-consumption model with exogenous interest rate $r$ and stochastic and exogenous income. The parallel is especially fitting because in this 2 period model, consumption in the second period is $c_1 = a_1(1 + r) + \iota_1$, which is linear in wealth $a_1$ and the state-contingent income shock $\iota_1$ (the 2 state variables in this model).

In this 2-period example, savings ($a_1$) are determined from the following Euler equation:

\[
uc(t_0 - a_1) = \beta(1 + r)E[uc(a_1(1 + r) + \iota_1)].
\]

If $uc(\cdot)$ is convex (or $ucc(\cdot) > 0$), because of the Jensen’s inequality, a mean preserving
increase in the variance of $\iota_1$ leads to a higher $a_1$. Notice that $c_1 = a_1(1 + r) + \iota_1$ is linear in both $a_1$ and $\iota_1$ and yet there are precautionary savings. Thus, the key to generate precautionary savings is not in the nonlinearity of the policies (if anything, the linearized policies exaggerate the Jensen’s inequality if the true policy is concave. This is because the concavity of the policy offsets the convexity of the marginal utility of consumption, which gives rise to the Jensen’s inequality. Indeed, we slightly overstate precautionary behaviour in our numerical examples below). Instead what matters for capturing the Jensen’s inequality is that, owing to the future policy function, we solve the stochastic version of the equilibrium conditions $(Ef(x', y', x, y) = 0)$, rather than their certainty equivalence counterpart $(f(Ex', Ey', x, y) = 0)$.

Finally, notice that the derivation of the policies from $\hat{E}_\tau f(x', g_{\hat{E}}(x'), \bar{x}_\tau, y) = 0$ is still a straightforward application of the implicit function theorem because, even with the expectations operator, we still have as many equations as unknowns.

### 3.4 Euler errors and accuracy

A solution with our algorithm has the following 2 properties, up to a numerical error

(a) At each $t$, $f(x_{t+1}, g_{\hat{E}}(x_{t+1}), x_t, y_t) = 0$ ($\hat{E}_t f(x_{t+1}, g_{\hat{E}}(x_{t+1}), x_t, y_t) = 0$ if one solves the state contingent version of the model).

(b) Convergence to a steady state.

Property (a) follows from step 4 in our algorithm and implies that we obtain very small Euler errors (up to numerical precision). That our algorithm guarantees very small Euler errors on the computed equilibrium path is reassuring. However, it should be noted that properties (a) and (b) are necessary but not sufficient for an accurate solution.

To see this assume a deterministic model. A true solution $\{x^*_t, y^*_t\}$ is such that

(a*) $f(x^*_{t+1}, y^*_{t+1}, x^*_t, y^*_t) = 0$ at all $t$.

(b*) Transversality conditions are not violated.

Does our solution satisfy these conditions, i.e. do (a) and (b) imply (a*) and (b*)? Property (b) (convergence to a steady state) rules out violation of transversality condi-
tions. However, property (a) does not imply \((a^*)\). The reason is that function \(g_{t+1}\) may not be equal to true policy function \(g\) and as a result a candidate solution \((x_{t+1}, g_{t+1}(x_{t+1}), x_t^*, y_t)\) that solves (a) would differ from \((x_t^*, y_t^*, x_t^*, y_t^*)\) where \(y_{t+1}^* = g(x_{t+1}^*)\).

To summarize, we are guaranteed to have very small Euler errors with our algorithm, but strictly speaking this does not guarantee an accurate solution.

While, as with any numerical method, we cannot guarantee that we are close to a true solution in general, below we prove that our algorithm performs better than the standard perturbation case in the neoclassical growth model.

3.5 Theoretical results in the neoclassical model

In this section we show that the solution attained with our one-step algorithm welfare dominates the one attained with the steady state linearization approach.\(^{23}\)

Since we compare to the perturbation method around the steady state, which assumes certainty equivalence, for a sharp comparison, here we focus on a deterministic problem.

Consider the value function associated with the Neoclassical model that has been introduced in section 3.2:

\[
v(k_t) = \max_{c_t, k_{t+1}} u(c_t) + \beta v(k_{t+1}) \quad \text{s.t.} \quad k_{t+1} = Ak_t^\alpha + (1 - \delta)k_t - c_t, \quad (8)
\]

Associated with the policy function for consumption linearized around the steady state, \(g_{ss}\), there is the following value function:

\[
v_{g_{ss}}(k_t) = u(g_{ss}(k_t)) + \beta v_{g_{ss}}(k_{t+1}) \quad \text{where} \quad k_{t+1} = Ak_t^\alpha + k_t(1 - \delta) - g_{ss}(k_t). \quad (9)
\]

The notation \(v_g\) highlights that this value function depends on the policy \(g\).

The following lemma finds the value function that is associated with the solution to our one-step algorithm.

\(^{23}\)The justification for assessing our solution through welfare is that the neoclassical model is efficient so the true solution achieves the highest welfare.
Lemma 1. Our one-step approach applied to the neoclassical model (8) is equivalent to solving the following problem:

\[
Tv_{g_{ss}}(k_t) = \max_{c_t,k_{t+1}} u(c_t) + \beta v_{g_{ss}}(k_{t+1}) \quad \text{s.t.} \quad k_{t+1} = Ak_t^{\alpha} + k_t(1 - \delta) - c_t. 
\]

Proof. To show that the Euler equation that we solve with our approach characterizes the solution to problem (10), start by taking the FOC from this problem w.r.t. \( c_t \). We get:

\[
u'_{g_{ss}}(k_t) - \beta \nu'_{g_{ss}}(k_{t+1}) = 0.\]

The envelope condition is:

\[
\nu'_{g_{ss}}(k_t) = u'(c_t)g'_{ss}(k_t) + \beta \nu'_{g_{ss}}(k_{t+1})(\alpha Ak_t^{\alpha-1} + (1 - \delta) - g'_{ss}(k_t)) = \beta u'(c_t)(\alpha Ak_t^{\alpha-1} + 1 - \delta) = u'(c_t)(\alpha Ak_t^{\alpha-1} + 1 - \delta).\]

Combining with the FOC and our next period policy \( c_{t+1} = g_{ss}(k_{t+1}) \), we get:

\[
u'(c_t) = \beta u'(g_{ss}(k_{t+1}))(\alpha Ak_{t+1}^{\alpha-1} + 1 - \delta),\]

which is the equation we solve (along with the resource constraint) in step (4) with our approach. It is easy to show that \( v_{g_{ss}} \) is concave, so the Euler equation is also sufficient.

We want to know whether our one-step approach is preferable to just using \( g_{ss} \) at \( k \) different from the steady state level \( k_{ss} \), i.e. we want to know whether the associated value \( Tv_{g_{ss}}(k) \) is closer to the true one \( v(k) \) than \( v_{g_{ss}}(k) \). This is indeed the case as affirmed in the next proposition.

Proposition 1. \( v(k) \geq Tv_{g_{ss}}(k) \geq v_{g_{ss}}(k) \) for any \( k > 0 \).

Proof. The first step consists of showing that \( Tv_{g_{ss}}(k) \geq v_{g_{ss}}(k) \). Let \( Tg_{ss} \) be the solution to (10). From (10), the welfare of using \( Tg_{ss} \) in an initial period at \( k \) and \( g_{ss} \) thereafter is \( Tv_{g_{ss}}(k) \). Instead the welfare of using \( g_{ss} \) in both the initial period and thereafter is \( v_{g_{ss}}(k) \). Assume by contradiction that \( Tv_{g_{ss}}(k) < v_{g_{ss}}(k) \). Then policies \( g_{ss} \) would give higher welfare than the new policy \( Tg_{ss} \). But then \( Tg_{ss}(k) \) would not be solving (10).

As the next step, let us show that \( Tv_{g_{ss}}(k) \leq v(k) \). This is because problems (8) and (10) only differ in the continuation value function and \( v(k) \geq v_{g_{ss}}(k) \) given that \( v \) is associated with the optimal policy.

Combining the inequalities in steps 1 and 2, \( v(k) \geq Tv_{g_{ss}}(k) \geq v_{g_{ss}}(k) \).

\[\square\]
The proposition is related to the analysis of the convergence properties of the policy iteration algorithm. Puterman & Brumelle (1979) and Santos & Rust (2004) among others have shown that, under certain conditions, the algorithm induces a monotonic improvement in the sense that at each iteration $Tv \geq v$. The two algorithms differ in many ways: we are not solving for a fixed point in the policy space and we are only solving the policy locally. Yet there is a relationship with a key step of the policy iteration algorithm: to solve at each point $x$, the policy iteration algorithm finds the new choice variables using the past iteration of the policy, evaluated at $x'$, the continuation point from $x$. This step is crucial for the result in the proposition above and is a commonality between the two algorithms.

That we can trace the result to the policy iteration algorithm is reassuring because, even though the results are only proven under somewhat restrictive conditions, in practice that algorithm is found to be rather robust to different models.

The proposition above covers the case in which we take only 1 backward step ($T_{aux} = 1$) from $k_{ss}$ to $k$. In section 3.3 we described the algorithm allowing for further intermediate backward steps ($T_{aux} > 1$). This way, at $k$ the continuation policy is not the steady state one but a policy $\hat{g}$ approximated at some point closer to $k$. Intuitively, this improves the accuracy even further. However, for the proof to go through, we need to guarantee that the continuation value that replaces $v_{ss}(k_{t+1})$ in (10), $v_{\hat{g}}(k_{t+1})$, is greater than $v_{ss}(k_{t+1})$. While intuitive, this is not necessarily the case. We can solve this issue by selecting the continuation policy that gives the highest welfare among the set $\{\hat{g}\}$ that collects all the policies computed on the backward loop, including the steady state one. This can be done as follows:

$$Tv_{\hat{g}}(k_t) = \max_{c_t, k_{t+1}, \{\hat{g}\}} u(c_t) + \beta v_{\hat{g}}(k_{t+1}) \quad \text{s.t.} \quad k_{t+1} = \alpha k_t^\alpha + k_t(1 - \delta) - c_t,$$

where $\{\hat{g}\}$ contains all the previously computed approximations backward from the steady
Since maximizing over \( \{ \hat{g} \} \) is computationally costly, in practice we guess that the best future policy is the previously computed one on the backward path. Later we confirm numerically that the multistep algorithm is indeed more accurate than the one with only one step, but more time consuming; see table 1. Nonetheless, the numerical application shows that even the one step algorithm is fairly accurate. Our theoretical result in proposition 1 is also related to the results reported in Bertsekas (2021), who find that even a single online step of the policy iteration algorithm often leads to substantial improvements in numerical accuracy.

4 Multicountry RBC model with capital non-reversibility and adjustment costs

We consider the multicountry RBC model with capital adjustment costs and capital irreversibility from Brumm & Scheidegger (2017). Each of the \( N \) countries is populated by an infinitely-lived representative agent with CRRA utility. Capital adjustment costs make the state space non-trivial: it is not enough to know the total amount of resources in the aggregate resource constraint but one needs to keep track of capital levels in each country as separate state variables. Capital irreversibility is an occasionally binding constraint which makes this model highly nonlinear. The fact that the model is multicountry makes it scalable and enables us to assess the ability to deal with a large number of state variables.

The aggregate resource constraint is:

\[
\sum_{i=1}^{N} (c_{i,t} + k_{i,t+1} + \Gamma(k_{i,t}, k_{i,t+1}) - k_{i,t}(1 - \delta) - A_{i,t}k_{i,t}^{\alpha}) = 0
\]

where

\[
\Gamma(k_{i,t}, k_{i,t+1}) = \frac{\phi}{2} k_{i,t} \left( \frac{k_{i,t+1}}{k_{i,t}} - 1 \right)^2
\]

are the capital adjustment costs.
We also require that the investment in each country $i$ remains non-negative:

$$k_{i,t+1} - k_{i,t}(1 - \delta) \geq 0, \quad \forall t.$$ 

### 4.1 Accuracy in a one country model

We start by solving this model with a single country, $N = 1$. With only one country we can get a global solution which we trust to be a very good approximation to the true one, and to which we compare the one from our algorithm. For $N > 1$, it becomes hard to get a solution that is accurate to the point of being considered the “true” solution. So, to assess the accuracy we mainly focus on $N = 1$. Instead, the case with many countries is mainly to illustrate how our algorithm addresses the curse of dimensionality.

To obtain the global solution to the one country model we adopt a policy iteration algorithm that uses cubic splines to approximate future policy functions. We use 2001 grid points for capital and approximate the AR(1) process for the TFP shocks with a Markov chain using the Rouwenhurst method with 17 states.

We start by setting parameter values as in Brumm & Scheidegger (2017): $\beta = 0.99$, $\delta = 0.01$, persistence of TFP shocks $\rho = 0.99$, the standard deviation of the shock innovations $\sigma_e = 0.01$, CRRA parameter $\gamma = 2$ and capital adjustment cost parameter $\phi = 0.5$. The global solution and our dynamic perturbation gives virtually identical results. However, with the calibration above, the capital irreversibility constraint never binds because the investment remains strictly positive.\(^{24}\) To obtain a simulation where the capital irreversibility constraint binds, we set $\beta = 0.96$, $\delta = 0.007$, $\rho = 0.8$, and $\sigma_e = 0.017$.

Figure 2 compares the simulated paths of length $T = 1000$ from our certainty equivalence dynamic perturbation method to the one from the accurate global solution.\(^{25}\) As one can

\(^{24}\)Brumm & Scheidegger (2017) report the results for $N \geq 2$ counties. We have found that increasing the number of countries makes it more likely that the capital irreversibility constraint will be binding during the simulation.

\(^{25}\)We use the certainty equivalence version of our code because we do not expect the Jensen’s inequality to play a big role. In our next application, which features rare disaster shocks and asset prices, we use the version with expectations.
see from the plot, the capital irreversibility constraint binds (investment becomes equal
to 0) several times during the simulation (there are 3 spells between periods 350 and 500
and one around period 800). Once again, the simulated paths from both methods are
very close. The average absolute difference between the two simulated paths for capital is
0.0335 percent of output, while that for investment is 0.001 percent of output. To provide
some reference, this is about 40 times smaller than the distance between the global solution
and the one linearized around the steady state. Not surprisingly, the steady state solution
performs especially poorly during the episodes when the non-reversibility constraint binds,
as shown in figure 3.

In the previous exercise, we start the simulation at the steady state, and as a result,
the simulated paths remain fairly close to the steady state. Since one of the advantages of
our dynamic perturbation algorithm is its ability to accurately compute transition paths
that start far away from the steady state, figure 4 shows what happens when we start the
simulation with only 60 percent of the steady state capital: the steady state solution leads
to large errors in the initial periods of transition whereas our solution remains uniformly
accurate even far away from the steady state.

We have some leeway in choosing certain parameters in our algorithm that can make
it run faster with somewhat less accurate results, or increase the accuracy of the results
at the cost of longer running time. As we discussed in Section 3, one of these parameters
is the length of the auxiliary path. The results above were obtained by setting it to $T_{aux} = 30$ as we found negligible accuracy improvements with a longer auxiliary path.
Table 1 shows how the accuracy and running time change as we change $T_{aux}$. For instance,
we can reduce the running time by a factor of 3 if we set $T_{aux} = 10$, at the cost of increasing
the average distance from the global solution by a factor of 2. Our fastest version, the one
step algorithm ($T_{aux} = 1$) takes 1.02 seconds to run. In terms of accuracy, the average
absolute difference between the simulated paths for capital computed with the global
solution and our algorithm with $T_{aux} = 1$ is 0.1050% of output. The corresponding error
Figure 2: Simulated paths for capital and investment from dynamic perturbation and global solution

Figure 3: Difference between simulated paths from dynamic perturbation and global solution, and linearisation around the steady state (SS) and global solution (as percentage of output from global solution)
with steady state policies is 1.2825%; so even the one step version of our algorithm makes the solution about 12 times more accurate than the steady state one (as mentioned, the 30 step solution is 40 times better). These results are in line with the findings of Bertsekas (2021) who argues that usually even only a one time online step can make large improvements over the “offline” solution (which in our case is the steady state policy).

Notice that these exercises are performed with a simulation of $T = 1000$ periods. As shown in Table 3, computing time is essentially proportional to $T$.

There are many other ways in which one could speed up the code: for instance, in non reported results we have also experimented with updating the policy functions along the auxiliary path only when the change in the state variables is sufficiently large.

4.2 Many countries

Next, we show how our algorithm performs in the model with many countries (and thus many state variables). The main purpose is to show how our method deals with the curse of dimensionality. All our results were obtained using Matlab R2021a on a laptop with

Figure 4: Difference between simulated paths from dynamic perturbation and global solution, and linearisation around the steady state (SS) and global solution (as percentage of output from global solution) when $k_0 = 0.6k$
Table 1: Accuracy and speed in a model with capital irreversibility with $N = 1$

| $T_{aux}$ | mean($\frac{|K_{DP}-K_{Glob}|}{Y_{Glob}} \times 100$) | mean($\frac{|I_{DP}-I_{Glob}|}{Y_{Glob}} \times 100$) | CPU (seconds) |
|-----------|---------------------------------|---------------------------------|--------------|
| 1 (one step) | 0.105                           | 0.008                           | 1.02         |
| 10        | 0.080                           | 0.003                           | 5.67         |
| 30        | 0.034                           | 0.001                           | 17.21        |
| 50        | 0.038                           | 0.0009                          | 26.73        |

an Intel(R) Core(TM) i9-9980HK CPU @ 2.40GHz processor and 32Gb of RAM.

The last column in table 2 reports the time (in seconds) it takes to generate a simulated path of length $T = 1000$ with our certainty equivalence dynamic perturbation method for models with $N = 1$, $N = 4$, $N = 50$ and $N = 100$ countries. The time increases in $N$. However, the increase is far less than the exponential rate typical of standard global methods. In particular, when using a numerical method that approximates the state space with tensor product grid, the number of grid points, and thus the time it takes to compute the model, grow exponentially with the number of state variables. This makes it unfeasible to solve the model with more than a few countries. Recent advances in computational methods substantially lessened the curse of dimensionality by cleverly distributing the grid points over the state space but this is typically successful in smooth models. Brumm & Scheidegger (2017) are the first ones that try this approach in a model with capital irreversibility. However, they report the results for only up to $N = 10$ countries. They solve for $N = 50$ countries in the smooth model without the irreversibility constraint.

The smooth version of the multicountry RBC model (without capital irreversibility) has become a standard benchmark to assess numerical methods aimed at dealing with the curse of dimensionality. See the special issue of the Journal of Economic Dynamics and Control (Den Haan et al. (2011)). More recently, Maliar & Maliar (2015) use that model to test a cutting edge solution method specifically designed to deal with a large number of state variables (200 countries, i.e. 400 state variables). In Appendix B we compare our results to theirs. The takeaway is that we have also been able to solve the model with 200 countries. Furthermore, the running times to generate equilibrium paths of length typical
of macro data with our solution method are substantially smaller than those reported in Maliar & Maliar (2015).

It should be noted that, in our algorithm, computation time increases with the number of periods in the simulation. This is shown in Table 3: using the model with \( N = 4 \) countries, we obtain the simulated paths of lengths \( T = 100, T = 500 \) and \( T = 1000 \). As one can see from the table, the time it takes to compute the path increases essentially linearly with the number of periods, \( T \). As a result, the algorithm in Maliar & Maliar (2015) or Brumm & Scheidegger (2017) that provide global policy functions would be more practical if one is interested to repeatedly simulate the model for a given set of parameter values. However, in many applications, such as model estimation, one needs to run the algorithm for many different values of the parameters but only once for each parameter specification. Our algorithm appears to be more suitable for this task.

Table 2: Accuracy and speed in multi-country model with capital irreversibility

<table>
<thead>
<tr>
<th>( N )</th>
<th>( L_1 )</th>
<th>( L_\infty )</th>
<th>CPU (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-5.18</td>
<td>-3.99</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>-5.26</td>
<td>-2.17</td>
<td>36</td>
</tr>
<tr>
<td>50</td>
<td>-4.53</td>
<td>-2.80</td>
<td>516</td>
</tr>
<tr>
<td>100</td>
<td>-4.60</td>
<td>-2.96</td>
<td>2320</td>
</tr>
</tbody>
</table>

Note: \( L_1 \) and \( L_\infty \) are the average and maximum Euler errors along our simulated path.

In table 2, we also report the average (\( L_1 \)) and maximum (\( L_\infty \)) Euler errors along our simulated path, which are the standard measures of accuracy used in the literature. Notice that our algorithm minimizes these errors by construction, see step (4) in our algorithm. However, since in this section we are using the certainty-equivalence version of our algorithm (we solve for \( f(x, y, x', y') = 0 \) along the equilibrium path), we report the Euler errors that embed expectations \( E[f(x, y, x', y')] = 0 \) to test for the imprecision due to not accounting for the Jensen’s inequality.\(^{26}\)

\(^{26}\)For models with \( N = 1 \) and \( N = 4 \), we approximate expectations using 5-point quadrature rules, while for \( N = 50 \) and \( N = 100 \), we used monomial rules described in Maliar & Maliar (2014) (since more accurate quadrature approximations become infeasible).
Table 3: Speed with different simulation length, model with capital irreversibility, \( N = 4 \)

<table>
<thead>
<tr>
<th>( T )</th>
<th>CPU (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>4</td>
</tr>
<tr>
<td>500</td>
<td>18</td>
</tr>
<tr>
<td>1000</td>
<td>36</td>
</tr>
</tbody>
</table>

5 Solving models with rare disaster shocks and zero-lower bound

In this section, we use our algorithm to solve a model with rare disasters from Fernandez-Villaverde & Levintal (2018). Models with rare disaster risk have been successfully used to account for many asset pricing puzzles, including the equity premium and risk-free rate puzzles.\(^{27}\) As argued by Fernandez-Villaverde & Levintal (2018), the mechanism that allows one to address the asset pricing puzzles in the models with rare disasters is the “large precautionary behaviour responses that are induced in normal times by the probability of the tail events”. To emphasise the effects of disaster risk, they use recursive (Epstein-Zin) preferences with production, capital adjustment costs, and nominal rigidities. They argue that this is an especially challenging model to compute numerically and that the standard perturbation method around the steady state (even of high order) is unable to fully capture the effects of the precautionary behaviour on asset prices. Instead, they use their innovative “Taylor projection” algorithm. They test the accuracy of their method by comparing the economic outcomes to those from a state-of-the-art global solution method (using Smolyak collocation method to economize on grid points given that there are 12 state variables). We make the same comparisons. In particular, we show that, using our algorithm without assuming certainty equivalence, we can replicate their results, obtaining a low risk-free rate and high equity premium. This illustrates the ability of our algorithm to capture the effects of expectations and precautionary behaviour on

\(^{27}\)See, for instance, Barro (2006), Gabaix (2011) and the list of references in Fernandez-Villaverde & Levintal (2018).
In addition, we expand the model of Fernandez-Villaverde & Levintal (2018) with one new feature that makes it even harder to compute: the zero lower bound (ZLB) on the nominal interest rate. This feature has recently received a lot of attention. Among others, Guerrieri & Iacoviello (2015), Maliar & Maliar (2015), Fernandez-Villaverde et al. (2015) and Aruoba et al. (2018) use different cutting-edge numerical methods to solve models with ZLB, but without capital and disaster shocks. We are not aware of any paper that has combined both rare disasters and the ZLB. The additional challenge that the ZLB poses for the numerical algorithm stems from possible kinks in the policy functions. Furthermore, to the extent that the ZLB does not bind in the steady state, it would not be possible to study the impact of the ZLB with solutions that, like Fernandez-Villaverde & Levintal (2018) and standard perturbation methods, rely on solving the model around the steady state.

We show that the addition of the ZLB in conjunction with disaster shocks has some interesting consequences. In particular, we have noticed that with the benchmark calibration of Fernandez-Villaverde & Levintal (2018) the nominal interest rate is negative. Imposing the ZLB increases the risk free rate and reduces the equity premium. In addition, when the ZLB binds, the effects of disaster risk are reversed: the higher the disaster risk, the lower the equity premium. These results highlight additional challenges that models with nominal frictions pose for the disaster shock mechanism to address the equity premium and risk-free interest rate puzzles.

5.1 Model setup

We consider the model from Fernandez-Villaverde & Levintal (2018), which is a New Keynesian model to which they add recursive preferences and rare disaster shocks that lead to some depreciation of capital, following Gurio (2012). Furthermore, the size of

\footnote{Numerically, we use the same approach to approximate expectations as Fernandez-Villaverde & Levintal (2018): by using quadrature or monomial rules. As a result, we manage to capture the same effects of precautionary behaviour with our algorithm.}
the disaster is time-varying and follows an AR(1) process. See Fernandez-Villaverde & Levintal (2018) for details.

5.2 Results

To solve this model, we use the version of our algorithm with expectations. As argued by Fernandez-Villaverde & Levintal (2018), rare disasters generate a strong impact on asset prices and risk premia, and the solution methods should be able to approximate these effects. Therefore, to assess their method, they compare average asset levels (capital in this model) and prices (equity and risk-free returns) obtained with their Taylor projection method to those from a state of the art Smolyak collocation method. Furthermore, they also report results from perturbation methods around the steady state, up to the 5th order. Table 4 reports these numbers from their paper, as well as the results from our algorithm. It is worth to highlight that, as shown below, the key to generate high equity premia in this set-up through disaster risk is in the ability to generate a low risk-free rate. Of the many methods that Fernandez-Villaverde & Levintal (2018) have tried, their two most accurate ones in terms of Euler errors (Taylor Projection and Smolyak Collocation of order 3) give a mean risk-free rate of 1.5-1.6%. Ours lies in between (1.55%). By comparison, they find that perturbation solutions around the steady state generate a much higher risk-free rate, ranging from 4.6% at the first order to 2.1% at the fifth order. The inability to match low risk-free rate also implies that the equity premium is lower with perturbation methods. Instead, our method produces an equity premium which is very similar, in fact slightly higher than the one with their global method (ours is 4.2% while theirs is 3.9%). The large equity premium is also reflected in the average level of capital and investment, slightly lower than the one with the global method (average capital is 1.8761 with our method and 1.8838 with theirs). Notice that in this model, the higher

\footnote{Our Euler errors are smaller, but we do not report them because, as previously discussed, they are minimized by our algorithm. Moreover, in this section we are using the algorithm with expectations so they would not even serve as a measure of the importance of the Jensen’s inequality as done in section 4.2.}
the risk of disaster, the lower capital and investment, as shown further below. That, if anything, our method overstates precautionary behaviour is consistent with the analysis in section 3 and it is expanded upon in appendix C by means of the simple RBC example. Since we are able to solve that model with an accurate global solution, we can quantify more precisely the extent to which our algorithm overstates precautionary savings. That appendix confirms that there is an overstatement of precautionary behaviour. However, it is minor, as one can appreciate from Figure 11.

Table 4: Simulation averages from different solution methods

<table>
<thead>
<tr>
<th>Method</th>
<th>$R_f$</th>
<th>$R_e - R_f$</th>
<th>$K$</th>
<th>$I$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perturbation (1st)</td>
<td>4.6</td>
<td>0.4</td>
<td>1.9695</td>
<td>0.0563</td>
</tr>
<tr>
<td>Perturbation (5th)</td>
<td>2.1</td>
<td>3.6</td>
<td>1.8835</td>
<td>0.0536</td>
</tr>
<tr>
<td>Taylor Projection</td>
<td>1.6</td>
<td>3.8</td>
<td>1.8849</td>
<td>0.0537</td>
</tr>
<tr>
<td>Smolyak Collocation</td>
<td>1.5</td>
<td>3.9</td>
<td>1.8838</td>
<td>0.0536</td>
</tr>
<tr>
<td>Dynamic Perturbation</td>
<td>1.6</td>
<td>4.2</td>
<td>1.8761</td>
<td>0.0534</td>
</tr>
</tbody>
</table>

Table 5 shows the results we obtain by solving the model for different values of the disaster probability ($p_d$). In line with the results in Fernandez-Villaverde & Levintal (2018) and with the usual precautionary behaviour intuition, through simulations, we find that as $p_d$ increases, average investment ($I$) and the size of the capital stock ($K$) decrease. As a result, the return on equity ($R_e$) increases. The rate of return on the safe real bond ($R_f$) drops rather fast. The decline in $R_f$ is much larger than the increase in $R_e$ and so the former mainly accounts for the large equity premium in this model. So, there is a level of disaster risk through which the model explains these asset prices facts. However, as we discuss next, the nominal predictions of this model uncover some counterfactual results implied by this approach.

The last column of table 5 shows that the nominal interest rate ($R_{nom}$) is decreasing in the disaster probability $p_d$, and the net nominal interest rate $R_{nom} - 1$ becomes negative for sufficiently high $p_d$ (the cut-off level for which $R_{nom} - 1 = 0$ is $p_d = 0.0024$ and generates an equity premium of 2.5% and a risk free rate of 3%). So the ability to generate the large equity premium and low risk free rate comes at the cost of having a negative nominal
interest rate.

Table 5: Simulation averages for different levels of disaster risk without ZLB

<table>
<thead>
<tr>
<th>$p_d$</th>
<th>$R_e$</th>
<th>$R_f$</th>
<th>$K$</th>
<th>$I$</th>
<th>$R_e - R_f$</th>
<th>$\pi$</th>
<th>$R_{nom}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000</td>
<td>1.0536</td>
<td>1.0525</td>
<td>2.0075</td>
<td>0.0571</td>
<td>0.0011</td>
<td>1.0041</td>
<td>1.0170</td>
</tr>
<tr>
<td>0.0010</td>
<td>1.0546</td>
<td>1.0438</td>
<td>1.9784</td>
<td>0.0563</td>
<td>0.0108</td>
<td>0.9989</td>
<td>1.0097</td>
</tr>
<tr>
<td>0.0020</td>
<td>1.0555</td>
<td>1.0351</td>
<td>1.948</td>
<td>0.0554</td>
<td>0.0203</td>
<td>0.9937</td>
<td>1.0024</td>
</tr>
<tr>
<td>0.0030</td>
<td>1.0563</td>
<td>1.0265</td>
<td>1.9168</td>
<td>0.0546</td>
<td>0.0298</td>
<td>0.9885</td>
<td>0.9952</td>
</tr>
<tr>
<td>0.0040</td>
<td>1.0571</td>
<td>1.0179</td>
<td>1.8850</td>
<td>0.0537</td>
<td>0.0392</td>
<td>0.9834</td>
<td>0.9881</td>
</tr>
<tr>
<td>0.0043</td>
<td>1.0573</td>
<td>1.0155</td>
<td>1.8761</td>
<td>0.0534</td>
<td>0.0418</td>
<td>0.9820</td>
<td>0.9861</td>
</tr>
</tbody>
</table>

5.3 Zero Lower Bound

Motivated by this observation, we add a new feature to the model of Fernandez-Villaverde & Levintal (2018): the constraint that the net nominal interest rate set by the monetary authority cannot be negative (the zero lower bound constraint). Figure 5 compares the predictions of the model with and without the ZLB, in terms of the real risk-free rate, nominal risk-free rate, inflation and equity premium.

The first thing to notice from this figure is that when $p_d$ is low, the ZLB does not bind ($R_{nom} > 1$ in panel 2), and both versions of the model produce essentially the same results: as $p_d$ increases, $R_f$, $R_{nom}$, and inflation decline, whereas the equity premium increases. This behaviour is not surprising.\(^{30}\) As we further increase $p_d$, the nominal rate continues to decrease in the model without the ZLB (dashed line in panel 2). In the model with the ZLB, on the other hand, the nominal interest rate remains stuck at its lower bound ($R_{nom} = 1$ in the figure). Since in this version of the model, the nominal rate cannot respond to inflation any more, inflation falls much faster (see panel 3). In turn, from the Fisher equation, the rate of return on the real safe bond increases (panel 1) and so the equity premium declines (panel 4). So this model is not able to explain the equity premium and risk free rate puzzle when the ZLB is imposed.

\(^{30}\) $R_f$ decreases in $p_d$ through the precautionary motive emphasized by Fernandez-Villaverde & Levintal (2018): with more risk there is more demand for the safe asset, which drives its return down. That $R_{nom}$ comoves with $R_f$ is consistent with the Fisher equation given that inflation also comoves with $R_{nom}$. And the comovement of $R_{nom}$ and inflation is a byproduct of the Taylor rule.
To summarize, these results highlight additional challenges that models with nominal frictions pose for the disaster shock mechanism to address the asset pricing puzzles. Furthermore, whether the response in times when the ZLB binds is consistent with the data is an open question which seems worth investigating given the attention that both the ZLB and disaster risk received in the literature. But this is left for future research.

6 An application to a model with sudden stops: accuracy and estimation with non-linear Kalman filter

The purpose of this section is to illustrate how our method can be combined with a non-linear time varying Kalman filter (KF) for model estimation and recover the true
parameters more accurately than the widely used standard KF that uses model linearization around the steady state. As a laboratory model for this section we use the seminal model of sudden stops in emerging economies of Mendoza (2010). In this environment a sudden stop is an episode when the borrowing limit that the small open economy faces becomes binding. This is an especially useful application to highlight the merits of our approach because the borrowing constraint does not bind in steady state but is “occasionally binding” over the equilibrium path causing sudden stops. Because the constraint is not binding in the steady state, the parameters estimated with the steady state approach are ill-identified and in particular, the method is completely uninformative about a key parameter that characterizes the tightness of the borrowing limit – the model likelihood is flat with respect to such parameter. Instead, we show that our method combined with a time varying KF generates a model likelihood with a well-behaved peak near the true value of this parameter.

A further motivation is to show that our algorithm can deliver an accurate solution to models with occasionally binding borrowing constraints.

The model is a small open economy from Mendoza (2010) described in appendix D. The key equation is the following borrowing constraint:

$$q_t^b b_{t+1} - \phi R_t (w_t n_t + p_t \nu_t) \geq -\kappa q_t k_{t+1}$$

where $q_t^b$ is the price of the one-period intertemporal bond $b_{t+1}$. $\phi$ is the share of production costs for labour and an intermediate input $(w_t n_t + p_t \nu_t)$ that must be financed through intratemporal borrowings at the world interest rate $R_t$. $\kappa$ is the maximum share of the value of capital $q_t k_{t+1}$ that can be used as collateral. This parameter only appears in the borrowing constraint and it does not affect the policy functions linearized around the steady state given that the constraint is slack at the steady state.
6.1 Accuracy

We begin by showing the accuracy of our algorithm. To do that we compare our method with a global solution method based on policy iterations with a large number of grid points. We draw a path for the exogenous shocks — the world interest rate $R_t$ and TFP— to generate a simulated path for all endogenous variables. To assess accuracy, we compare the equilibrium variables obtained by the accurate global solution with the ones generated by our “dynamic perturbation” method when using the same exogenous shocks and initial conditions. Figure 6 shows the path obtained with the global solution, our dynamic perturbation with expectations, and the linearization around the steady state. From panel 4 it can be noticed how the borrowing limit binds in several occasions: the plot in this panel shows the residual in the borrowing limit

$$BL_t = q^h b_{t+1} - \phi R_t (w_t n_t + p_t \nu_t) + \kappa q_t k_{t+1}.$$ (11)

To satisfy the borrowing limit, we must have $BL_t \geq 0$. The borrowing limit is binding when $BL_t$ touches the horizontal line, i.e. $BL_t = 0$. Even in those episodes where the borrowing constraint is activated (and the policy functions are highly nonlinear), our simulated path remains accurate. In comparison, the steady state approach becomes especially inaccurate after the first episode in which the borrowing limit binds (period 16) in the global solution. As one can see from panel 4, $BL_t$ computed with the steady state solution is often negative which means that the borrowing limit is violated. The issue with the steady state approach is that the borrowing limit does not bind in the steady state and so any local approximation (linear or of higher order) to policies around the steady state ignore the effect of the constraint even when it should be activated. This will become especially important for the estimation through the Kalman Filter.

Also worth highlighting is the role of expectations. Figure 7 compares the same global solution to the certainty equivalence version of our algorithm. In this case, there is a
Figure 6: Simulated equilibrium paths of Mendoza (2010) using dynamic perturbation with expectations

$K$ is capital, $B$ bonds, $\text{AggC}$ is the utility aggregator $c_t - N(n_t)$ (see Appendix D), and $\text{BL}$ is the residual from the borrowing limit defined in (11).
larger difference with the global solution, especially for the path of bond positions and for how far we are from the borrowing limit. Intuitively, the precautionary motive to avoid becoming borrowing constrained is neglected with the certainty equivalence approach.

Table 6 reports the mean absolute deviation from the global solution. The numbers are expressed as a percentage of the utility aggregator $c - N(n)$ computed through the global method. The table confirms what can be seen from figures 6 and 7, i.e. that our solution is more accurate than the steady state and that the version with expectations improves accuracy further, especially for bonds, utility aggregator $c - N(n)$, and for the borrowing limit $BL$.

Figure 7: Simulated equilibrium paths of Mendoza (2010) using dynamic perturbation with certainty equivalence.

Note: $K$ is capital, $B$ bonds, AggC is the utility aggregator $c_t - N(n_t)$ (see Appendix D), and BL is the residual from the borrowing limit defined in (11).
Table 6: Mean absolute deviation from global solution as percentage of global utility aggregator

<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>B</th>
<th>AggC</th>
<th>BL</th>
<th>n</th>
<th>Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP (expectations)</td>
<td>0.219</td>
<td>0.529</td>
<td>0.095</td>
<td>0.477</td>
<td>0.040</td>
<td>0.057</td>
</tr>
<tr>
<td>DP (certainty equivalence)</td>
<td>0.226</td>
<td>3.010</td>
<td>0.249</td>
<td>2.691</td>
<td>0.048</td>
<td>0.076</td>
</tr>
<tr>
<td>Steady State linearization</td>
<td>0.452</td>
<td>8.992</td>
<td>0.700</td>
<td>8.122</td>
<td>0.061</td>
<td>0.079</td>
</tr>
</tbody>
</table>

Note: the errors are expressed as a percentage ratio of the utility aggregator $AggC = c - N(n)$ computed with the global solution.

6.2 The extended Kalman filter

One popular approach to evaluate the likelihood of the DSGE models is to combine the linear approximation of the model’s solution around the steady state with the standard linear KF. Intuitively, this approach may not perform well in highly non-linear models. Poor approximation of the model’s solution may lead to poor identification of the structural shocks by the KF and, as a result, poor estimation of the parameters.

Our solution method, on the other hand, along with a path of equilibrium state and jump variables, also generates a corresponding sequence of locally linear approximations to the policy and transition functions. We can use these local linear approximations to the policy and transition functions as inputs for the extended Kalman filter (EKF), the nonlinear version of the KF that linearizes the equations of the model around the best current estimate of the model’s state. The details of the EKF are provided in Havlik & Straka (2015) among others.\(^{31}\)

To illustrate how the EKF works, consider the following nonlinear state transition and observation model:

\[
x_{t+1} = h(x_t) + \eta_{t+1} \\
z_t = \phi(x_t) + \epsilon_t
\]

\(^{31}\)EKF was invented for the Apollo lunar mission and has become a standard in the theory of nonlinear state estimation and its applications such as satellite navigation systems, see Grewal & Andrews (2008).
where $x_t$ is the vector of the states and $z_t$ is the vector of observables (measurements). The first equation is the state transition equation, while the second one is the measurement equation, $\eta_t$ are structural shocks and $\epsilon_t$ are measurement errors. In comparison, with the standard KF, $h(x_t)$ and $\phi(x_t)$ are assumed to be linear functions. Let $\hat{x}_{t|t-1}$ and $\hat{x}_{t|t}$ be the best estimates (means) of the state vector in period $t$ given all the information up to period $t - 1$ and $t$ respectively. Given the measurement residual from period $t$:

$$e_t = z_t - \phi(\hat{x}_{t|t-1})$$

EKF updates the estimate of the state similarly to the regular KF:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t e_t$$

where $K_t$ is the Kalman gain which is computed as in the standard Kalman filter, but using the derivatives of $\phi(\cdot)$ and $h(\cdot)$ at the current best estimates of the state. See Durbin & Koopman (2012) for details.

The EKF can be naturally combined with our solution algorithm given that it generates local derivatives of $\phi(\cdot)$ and $h(\cdot)$. The procedure of mapping the linearized economic model into the appropriate elements of the Kalman filter model follows the standard steps in the literature (see for instance Fernandez-Villaverde et al. (2016)). The only difference is that our mapping from the policy functions to the KF matrices is done every period given the current estimate of the state.

6.3 Applying the extended Kalman filter to the model of Mendoza (2010)

How well can our approach with the EKF recover the true states, $x_t$? Recovering the states is a key step in standard model estimation where this procedure is embedded into a wider loop that maximizes the likelihood or posterior with respect to the model parameters, see for instance Fernandez-Villaverde et al. (2016). The success of this wider loop hinges on...
the reliability of the KF technique. To test our KF technique we will simulate data for \( x_t \) and \( z_t \) from our global solution and see if, given the observables \( z_t \), our procedure can recover the true states \( x_t \) behind the simulation.

We assume that \( z_t \), the vector of observables, consists of: (1) investment to output ratio \( \left( \frac{I_t}{Y_t} \right) \), (2) real interest rate \( (R_t) \), (3) hours worked \( (n_t) \), (4) output growth \( (\log(Y_t) - \log(Y_{t-1})) \), (5) current account to output ratio \( \left( \frac{CA_t}{Y_t} \right) \). Since there are only 2 structural shocks and 5 measurement equations, it is necessary to include measurement error \( \epsilon_t \) in 3 of the 5 measurement equations. We add small measurement errors (with standard deviation 0.001) in the first 3 observables above.

We assume that we have a precise estimate of the initial state and run both the standard linear KF that uses the model solution obtained by linearization around the steady state, and the EKF combined with our solution algorithm.\(^{32}\)

To get a sense of how well the EKF performs, we report the mean absolute percentage error between the true state variables and the ones estimated with the two versions of the Kalman filter.

<table>
<thead>
<tr>
<th></th>
<th>( K )</th>
<th>( B )</th>
<th>TFP</th>
<th>( R )</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>EKF/DP</td>
<td>0.0031</td>
<td>0.0025</td>
<td>0.0026</td>
<td>0.0026</td>
<td>0.0027</td>
</tr>
<tr>
<td>KF/SS</td>
<td>0.0040</td>
<td>0.0102</td>
<td>0.0034</td>
<td>0.0031</td>
<td>0.0052</td>
</tr>
</tbody>
</table>

Table 7: Mean absolute percentage errors from the Kalman filter

Note: in the first 4 columns we report the average error from 100 simulations of length \( T = 100 \) for each state variable (\( K, B, TFP, R \)). The last column is the average of the other columns.

As the table shows, the mean absolute percentage error is smaller with the EKF for all state variables. The most noticeable improvement is in terms of recovering the bonds \( b \), where the error is almost 4 times smaller than with the standard KF. On average, the error with the EKF is about half the one with the standard linear KF.

Next, we show the implications for the estimation of model parameters related to the

\(^{32}\)Assuming a precise estimate of the initial state (as well as very small measurement error \( \epsilon_t \)), we can attribute the discrepancy between the recovered states and the true ones to filter inaccuracy.
borrowing limit. The main advantage of our approach is its ability to identify \( \kappa \), the parameter that controls the tightness of the borrowing limit, which is not identified with the standard steady state approach. This is shown in the first panel of Figure 8 (dashed horizontal red line): the likelihood computed using the standard linear KF is entirely flat as a function of \( \kappa \). This result is not surprising since with the standard KF, the model solution is obtained by linearizing the model around the steady state, where the borrowing limit is slack; as a result, \( \kappa \) has no effect on the linear KF. On the other hand, the figure shows that the likelihood with our method (solid line) has a well-behaved peak close to the true value of the parameter (denoted by a dashed black vertical line). The right panel shows the likelihood as a function of \( \phi \), the other parameter of the credit constraint. In this case the likelihood with the linear KF is not flat. However, it keeps increasing in \( \phi \) until it hits the upper bound (0.28 in the figure) such that the borrowing limit starts to bind in the steady state.\(^{33}\) Our method instead has a peak close to the true value of \( \phi \). The reason why in this case the likelihood with the linear KF is not flat is that \( \phi \) also enters the budget constraint and thus it impacts the behaviour of the model around the steady state. However, the linear KF misses the implications of \( \phi \) for the credit constraint which explains why the estimation is far from the true parameter value.

\(^{33}\)We do not solve for higher \( \phi \) as we know that the borrowing limit does not bind in the steady state for the true parameters.
Figure 8: Model likelihoods as a function of ϕ and κ

Notes: other parameters are kept at their true value. The range of the plot is constrained by the fact that for κ too small and for ϕ too large, the borrowing limit binds at the steady state. We are not interested in solving for parameter values such that the borrowing limit binds in the steady state as we know that not to be the case for the true parameters. The plotted likelihood is the average over a 100 simulations of length $T = 100$.

7 Conclusion

In this paper, we develop a new algorithm which is suitable for solving large-scale nonlinear dynamic models over large transitions. The method consists of Taylor expanding the equilibrium conditions of the model not just around the steady state, but sequentially along the entire equilibrium path. We demonstrate that the method can be applied to a broad class of models. In particular, we show that it can solve models with a large number of state variables and strong non-linearities. We solved models with the typical examples of strong non-linearities in macroeconomics: due to capital irreversibility constraints, occasionally binding borrowing limits, rare disasters and zero lower bound. We also demonstrate how we can incorporate uncertainty and expectations into our algorithm and deal with the Jensen’s inequality.

Because our policies are locally linear, we can make use of an extended version of the Kalman filter with time varying coefficients to identify shocks from the data. With this tool in hand we are able to evaluate the likelihood function which can be used for the estimation of nonlinear models.
Besides the applications included in this paper, our method may be used to study the big events that are reshaping the world and involve either large shocks or large transitions such as global warming, decarbonization, pandemics, wars, unprecedented increases in public debt.\textsuperscript{34} It is our hope that our algorithm will be a useful addition to the toolkit of macroeconomists.

References


\textsuperscript{34}As an example, our methodology has been applied to study the growth and business cycle implications of large demographic transitions of the labor force by gender, age and education (Mennuni 2019).


A Applying the implicit function theorem

For simplicity, let’s consider the deterministic version of our algorithm.

After substituting next-period jump variables with a policy function \( y' = g_z(x') \), the equilibrium system of equations becomes:

\[
f(x', g_z(x'), x, y) = 0
\]  \hspace{1cm} (12)

which can be written as:

\[
G(x, z) = 0
\]

where \( z = (x', y) \) collects the variables we need to solve for, given the current value of the state vector \( x \). As explained in the main text, \( n_z = n_x + n_y \) is equal to the number of equations in \( G \). As long as the matrix of partial derivatives of \( G \) with respect to \( z \) evaluated at \( (x, z) \), \( G_z(x, z) \), is non-singular, we can use the Implicit function theorem to find:

\[
\frac{dz}{dx} = [G_z(x, z)]^{-1}G_x(x, z).
\]

We compute \( G_z(x, z) \) and \( G_x(x, z) \) using finite-difference approximations.

Note that the implicit function theorem cannot be applied with the usual steady state approach because without function \( g_z \), one has to solve for \( y' \) too and so there are more unknowns than equations.

For the algorithm with expectations, we only need to replace (12) with

\[
Ef(x', g_z(x'), x, y) = 0
\]
Again, there are as many equations as unknowns (see the discussion in section 3.3.1), so one can follow the same steps as above.

B Many countries without capital irreversibility

In this section we consider the multi-country RBC model adopted by Maliar & Maliar (2015), which does not have capital irreversibility as in section 4. Table 8 reports the summary of the results when using our “certainty equivalence” algorithm to generate a simulated path of length $T = 150$ for the case with $N = 20$, $N = 40$ and $N = 200$ countries. For this exercise, we set $\beta = 0.99$, $\delta = 0.025$, $\rho = 0.95$, $\sigma_\epsilon = 0.01$ and $\gamma = 1$ (logarithmic preferences). We set the initial capital $k_{i,0}$ equal to half their steady state levels for all countries. The initial TFP levels are set at the steady state. The table reports the average absolute size of the Euler equation errors ($L_1$) computed during the simulation, their corresponding maximal absolute value ($L_\infty$) and the running time in seconds (CPU) using Matlab on a standard desktop computer with one Intel core i7-6700 3.40GHz processor.\textsuperscript{35} We can compare these results with the ones reported in Maliar & Maliar (2015). The running times for our solution method are substantially smaller than those reported in Maliar & Maliar (2015): the fastest but the least accurate version of their algorithm that approximates policy functions using first-order polynomials takes 21 minutes with $N = 20$, 89 minutes with $N = 40$ and 105 minutes with $N = 200$. The Euler errors from our algorithm are also smaller than the ones reported in Maliar & Maliar (2015) for their fast version of the algorithm. The average Euler errors that they report when using their more precise (but significantly slower) version of the algorithm with second-order polynomial basis are slightly smaller than ours. However, unlike Maliar & Maliar (2015), we start our simulations far away from the steady state. We have used the code for their algorithm publicly available online, and obtained average Euler errors that are somewhat larger than from our algorithm when using the same initial conditions.

\textsuperscript{35}To obtain the Euler equation errors, we approximate expectations with 11-point quadrature.
Table 8: Accuracy and speed in multi-country model

<table>
<thead>
<tr>
<th>N</th>
<th>$L_1$</th>
<th>$L_\infty$</th>
<th>CPU (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>−5.43</td>
<td>−4.43</td>
<td>11</td>
</tr>
<tr>
<td>40</td>
<td>−5.42</td>
<td>−4.53</td>
<td>30</td>
</tr>
<tr>
<td>200</td>
<td>−5.42</td>
<td>−4.70</td>
<td>1085</td>
</tr>
</tbody>
</table>

and the same realizations of shocks.\(^{36}\)

An advantage of Maliar & Maliar (2015) is that it provides policy functions that can be repeatedly used to simulate equilibrium paths for alternative initial values of the state variables and random shock realizations.

### C Dealing with Expectations and Jensen’s Inequality

Section 5 finds that our algorithm with expectations slightly overstates precautionary behaviour. In the model of that section, precautionary behaviour leads to reduced investment. This is in contrast to the well known result that in neoclassical settings, precautionary behaviour leads to more capital accumulation. It is insightful to study precautionary behaviour in a simple neoclassical set up, a RBC model with CRRA preferences and see how results change when changing the intertemporal elasticity $\gamma$, and the variance of the shocks.

As mentioned, the deterministic version of our algorithm (that replaces future stochastic variables with their expected value) does not capture the impact of the size of the shocks on policy functions. In particular, it would not capture the precautionary savings behavior. However, as we have described in Section 3.3, one can incorporate expectations explicitly into our algorithm. In this section, we explore whether there are any benefits of doing so and to what extent the additional complexity makes our algorithm slower.

Intuitively, one may expect that taking expectations into account is more important

\(^{36}\)It should also be noted that the Euler errors generated by our solution method remain uniformly small during the whole simulated equilibrium path, while those generated by the solution method in Maliar & Maliar (2015) increase far away from the steady state.
when the shocks are larger, and when consumers are more risk averse. We consider two parameterizations. In both cases, we assume discount factor $\beta = 0.99$, capital share $\alpha = 0.36$, TFP persistence $\rho = 0.9$ and depreciation rate $\delta = 0.1$. In the first case, we set the variance of the shock $\sigma_e = 0.01$ and $\gamma = 2$, while in the second case, we set $\sigma_e = 0.05$ and $\gamma = 5$. To assess the accuracy of the solutions generated by the “deterministic” and “expectations-based” versions of our algorithm, we compare both of those to an accurate global solution obtained with a large number of grid points.\footnote{For the global solution we use policy function iteration with 500 gridpoints for capital with piecewise linear interpolation between grid points and Rouwenhorst discretization of the productivity shock with 25 gridpoints.} We draw a sequence of exogenous shocks and compare the simulated paths for capital generated by these alternative solutions.

Figures 9 and 10 show the results. The deviations (in absolute terms) from the simulated path generated by the accurate global solution is interpreted as the “error”. The two figures display these errors for the solution produced by our baseline algorithm with certainty equivalence (“Dyn. Perturb.”), our algorithm with expectations (“Dyn. Perturb.$_E$”) and by the standard linear approximation around the steady state.

Figure 9 shows that the addition of expectations into our algorithm produces a small
improvement on the solution errors in the setup with small volatility and low risk aversion. Figure 10, on the other hand, shows that the algorithm that incorporates the expectations produces substantially more accurate results in the setup with higher volatility of the shocks and higher risk aversion of consumers. At the same time, incorporating expectations (using a 5-point quadrature approximation) makes our algorithm about 40 percent slower.\footnote{We provide more detail about the algorithm speed in Appendix B}

To better illustrate the effect that the incorporation of expectations has on the solution, in Figure 11 we plot the simulated path for capital in the high-volatility / high-risk-aversion version of the model where in the simulation we set all TFP shock realizations to their average value (while there are no shocks in the particular simulation, agents expect shocks to occur). As one can see from that figure, the simulated path generated from our baseline algorithm, as well as the path generated by the first-order approximation around the steady state, display the certainty equivalence property because the simulations remain at the determinist steady state despite the presence of uncertainty in the model. Instead, our algorithm that incorporates expectations leads to similar precautionary savings to the global solution– the simulated paths for capital lie above the corresponding
Figure 11: Simulated path with no shocks in high-volatility / high-risk-aversion setup
deterministic steady state solution. In fact, we slightly overstate the amount of precau-
tionary savings. This might seem surprising because a common misperception is that
linear policies do not capture the Jensen’s inequality. In section 3.3.1 we have clarified
this point analytically by connecting it to the standard example of the precautionary
savings in a 2 period savings-consumption model.

D Sudden Stops model of Mendoza-2010

A representative consumer in a small open economy maximizes:

$$E_0 \left( \sum_{t=0}^{\infty} \exp \left( - \sum_{\tau=0}^{t} \rho (\bar{c}_\tau - N(\bar{n}_\tau)) \right) u(c_t - N(n_t)) \right)$$

subject to a sequence of the following constraints:

$$c_t + i_t + p_t \nu_t + \delta^t b_{t+1} = \exp(cA_t) F(k_t, n_t, \nu_t) - \phi(R_t - 1)(w_t n_t + p_t \nu_t) + b_t - \tau,$$

where

$$i_t = \delta k_t + (k_{t+1} - k_t) \left( 1 + \Psi \left( \frac{k_{t+1} - k_t}{k_t} \right) \right).$$
The notation follows Mendoza (2010). \( \tau \) is a lump sum tax used to finance an exogenous and constant stream of government spending.

Note that \( \rho(.) \) is an increasing and concave “endogenous discount rate” function. It solves the problem of continuum of deterministic steady states in a small open economy model, see Schmitt-Grohé & Uribe (2003). \( \bar{c} \) and \( \bar{n} \) denote the aggregate consumption and labor supply which consumers take as given (in equilibrium, \( \bar{c} = c \) and \( \bar{n} = n \)).

Output is produced using a constant-returns-to-scale technology that requires capital \( (k_t) \), labor \( (n_t) \), and an intermediate imported input \( (\nu_t) \). \( \varepsilon_t^A \) is a TFP shock. \( \phi \) is a fraction of the cost of labor that is paid in advance of sales with an intra-period working capital loan. International lenders charge the world interest rate \( R_t = R \exp(\varepsilon_t^R) \) on both the intra- and inter-period loans, where \( \varepsilon_t^R \) is the interest rate shock.

Additionally, domestic consumer faces the following collateral constraint:

\[
q_t^b b_{t+1} - \phi R_t (w_t n_t + p_t \nu_t) \geq -\kappa q_t k_{t+1}
\]

where \( q_t^b = 1/R_t, q_t = \partial i(\bar{k}_{t+1}, \bar{k}_t)/\partial \bar{k}_{t+1} \), and \( w_t = \partial N(\bar{n}_t)/\partial \bar{n}_t \). \( p_t \) is the price of foreign intermediate good \( \nu_t \) used in domestic production and for simplicity we assume it to be constant.\(^{39}\)

The functional forms of preferences and technology are as follows:

\[
u(c_t - N(n_t)) = \frac{(c_t - \frac{n_t^\omega}{\omega} + \xi)^{1-\sigma} - 1}{1 - \sigma}, \quad \sigma, \omega > 1, \xi > 0
\]
\[
\rho(c_t - N(n_t)) = \gamma \log \left( 1 + c_t - \frac{n_t^\omega}{\omega} + \xi \right), \quad 0 < \gamma \leq \sigma,
\]
\[
F(k_t, n_t, \nu_t) = A k_t^\alpha \nu_t^\eta, \quad 0 < \alpha, \beta, \eta < 1, \quad \alpha + \beta + \eta = 1, \quad A > 0,
\]
\[
\Psi \left( \frac{z_t}{k_t} \right) = \frac{a}{2} \left( \frac{z_t}{k_t} \right)^2, \quad a \geq 0.
\]

\(^{39}\)This way our global method, that we use to assess our dynamic perturbation algorithm, saves on one state variable and is more accurate. It would be easy to make \( \nu_t \) a stochastic variable in our dynamic perturbation algorithm.
The parameters should be self explanatory and their values come from Mendoza (2010).

\( \xi > 0 \) is such that \( c_t - \frac{ny}{\omega} + \xi \) > 0 all over the state space, which in turn guarantees that \( u(\cdot) \) is real valued.

\( \varepsilon^A_t, \varepsilon^R_t \) follow AR(1) processes.