Existence and uniqueness of solutions to dynamic models with occasionally binding constraints.

Tom D. Holden[1], School of Economics, University of Surrey

Abstract: We present the first necessary and sufficient conditions for there to be a unique perfect-foresight solution to an otherwise linear dynamic model with occasionally binding constraints, given a fixed terminal condition. We derive further conditions on the existence of a solution in such models. These results give determinacy conditions for models with occasionally binding constraints, much as Blanchard and Kahn (1980) did for linear models. In an application, we show that widely used New Keynesian models with endogenous states possess multiple perfect foresight equilibrium paths when there is a zero lower bound on nominal interest rates, even when agents believe that the central bank will eventually attain its long-run, positive inflation target. This illustrates that a credible long-run inflation target does not render the Taylor principle sufficient for determinacy in the presence of the zero lower bound. However, we show that price level targeting does restore determinacy in these situations.

Keywords: *occasionally binding constraints, zero lower bound, existence, uniqueness, price targeting, Taylor principle, linear complementarity problem*

JEL Classification: *C62, E3, E4, E5*

This version: 28 March 2016

The latest version of this paper may be downloaded from:
https://github.com/tholden/dynareOBC/raw/master/TheoryPaper.pdf

[1] Contact address: School of Economics, FASS, University of Surrey, Guildford, GU2 7XH, England
Contact email: thomas.holden@gmail.com Web-site: http://www.tholden.org/
The author would particularly like to thank Michael Paetz for joint work on an earlier algorithm for simulating models with occasionally binding constraints. Additionally, the author would like to thank the participants in numerous seminars at which versions of this paper were presented. Further special thanks are due to the following individuals for helpful discussions: Gary Anderson, Martin M. Andreasen, Paul Beaurdry, Saroj Bhattarai, Charles Brendon, Zsolt Csizmadia, Oliver de Groot, Michael B. Devereux, Bill Dupor, Roger Farmer, Michael Funke, William T. Gavin, Fabio Ghironi, Andy Glover, Luca Guerrieri, Pablo Guerrón-Quintana, Matteo Iacoviello, Tibor Illés, Michel Juillard, Hong Lan, Paul Levine, Jesper Lindé, Albert Marcet, Enrique Martinez-Garcia, Antonio Mele, Alexander Meyer-Gohde, Adrienn Nagy, Matthias Paustian, Jesse Perla, Frank Portier, Søren Ravn, Alexander Richter, Jonathan Swarbrick, Simon Wren-Lewis and Carlos Zarazaga. Financial support provided to the author by the ESRC and the EC is also greatly appreciated. Furthermore, the author gratefully acknowledges the use of the University of Surrey FASS cluster for the numerical computations performed in this paper, and the University of Washington for providing the office space in which the bulk of this paper was written.

1. Introduction

Since the financial crisis of 2007-2008, many central banks around the world have chosen to keep their nominal interest rate close to 0%. While in a few countries, rates on some assets have gone slightly negative, central banks are unable to push their target rate to the level a Taylor rule might suggest since agents always have the option of holding cash. In practice then, central banks face a zero lower bound (ZLB) on their policy rate, which limits their ability to provide stimulus in severe recessions. Furthermore, during the crisis, both households, firms and banks have hit their borrowing constraints, which has limited their ability to smooth out its effects. However, the theoretical results on determinacy which justify the Taylor principle do not apply to models with occasionally binding constraints (OBCs), such as the zero lower bound, or a borrowing constraint, meaning that the profession still lacks all of the necessary tools for understanding the behaviour of such models.

In this paper, we develop theoretical tools for understanding the behaviour of models with occasionally binding constraints.[2] Much as the seminal paper of Blanchard and Kahn (1980) provided necessary and sufficient conditions for the existence of a unique perfect foresight solution to a linear model that returns to a given steady-state, we will provide the first necessary and sufficient conditions for there to be a unique perfect foresight solution, returning to a given steady-state, in otherwise linear models with occasionally binding constraints. We will also provide both necessary conditions and sufficient conditions for there to exist any such solutions. When no solution returning to the given steady-state exists, this implies that the model must converge to some alternative steady-state.

As was observed by Benhabib, Schmitt-Grohé, and Uribe (2001a; 2001b), in the presence of OBCs, there are often multiple steady-states. For example, a model with a zero lower bound on nominal interest rates and Taylor rule monetary policy when away from the bound will have an additional "bad" deflationary steady-state in which nominal interest rates are zero. The presence of such multiple steady-states means that there can be sunspot equilibria which jump between the neighbourhoods of the two steady-states. Furthermore, if agents put a positive probability on being in the neighbourhood of the "bad" steady-state in future, then since this "bad" steady-state is indeterminate, then by a backwards induction argument, there is indeterminacy now. The consequences of

[2] A companion paper (Holden 2016), develops computational tools for understanding the same thing.

indeterminacy of these kinds have been explored by Schmitt-Grohé and Uribe (2012), Mertens and Ravn (2014) and Aruoba, Cuba-Borda, and Schorfheide (2014), amongst others. In all cases, the key to generating indeterminacy is that agents' beliefs about the point to which the economy would converge in the absence of future uncertainty are switching from one steady-state to the other.

However, the central banks of most major economies have announced inflation targets. Furthermore, it appears that in many countries, the central bank president would eventually be fired were their regime characterised by persistent deflation rather than the targeted (positive) inflation rate. This suggests that at least in the long-run, agents ought to believe that we will return to positive inflation, and they ought to place zero probability on paths converging to deflation. If agents' beliefs satisfy these restrictions, then the kinds of multiplicity studied by the authors cited in the previous paragraph are ruled out. It is an important question, then, whether there are still multiple equilibria even when all agents believe that in the long-run, the economy will return to a particular steady-state.

It is on such equilibria that we focus on in this paper, providing necessary and sufficient conditions for the existence of a unique perfect-foresight path, and also examining whether it is actually consistent with rationality for agents to believe that the economy will eventually return to the given steady state. A restricted class of such equilibria were also examined by Brendon, Paustian, and Yates (2015) who examined multiplicity in specific, non-standard, models when agents believe that with probability one, in one period's time, they will escape the bound and return to the neighbourhood of the "good" steady-state.

We show that many standard New Keynesian models featuring endogenous state variables (e.g. price dispersion), such as those of Fernández-Villaverde et al. (2012) or Smets and Wouters (2003; 2007) do not possess such a unique perfect-foresight path, meaning that even when agents' long-run expectations are pinned down, there is still multiplicity of equilibria. Thus, the Taylor principle is not sufficient for determinacy in the presence of occasionally binding constrains. Indeed, we show that in these models, there are some initial states from which the economy has one return path that never hits the ZLB, and another that does hit it, so the fact that the ZLB is not violated in a model in which it is not imposed does not mean that it would not be hit were it to be imposed. However, we show that under a price-targeting regime, there is a unique equilibrium path even when we impose the ZLB. Thus, if policy makers were convinced by the arguments for the Taylor principle, then, given they face the zero lower bound, they ought to consider adopting a price level target.

We also provide both necessary and sufficient conditions for the existence of any perfect-foresight solutions which return to the original ("good") steady-state. When no such equilibrium exists, agents must switch their beliefs to the other ("bad") steady-state, where they will remain in the absence of any way for agents to coordinate back on the "good" steady-state. We show that for standard New Keynesian models with endogenous state variables, there is a positive probability of ending up in a state of the world in which there is no perfect foresight path returning to the "good" steady-state,[3] implying that in the stochastic model, agents must always put positive probability on tending to the "bad" steady-state. This in turn implies global indeterminacy in such models, by a backwards induction argument. Once again though, price level targeting is sufficient to restore determinacy.

The key idea behind all of our proofs is that an OBC provides a source of endogenous news about the future. When a shock hits, driving the economy to the bound in some future periods, that tells us that in those future periods, the (lower) bounded variable will be higher that it would be otherwise.[4]

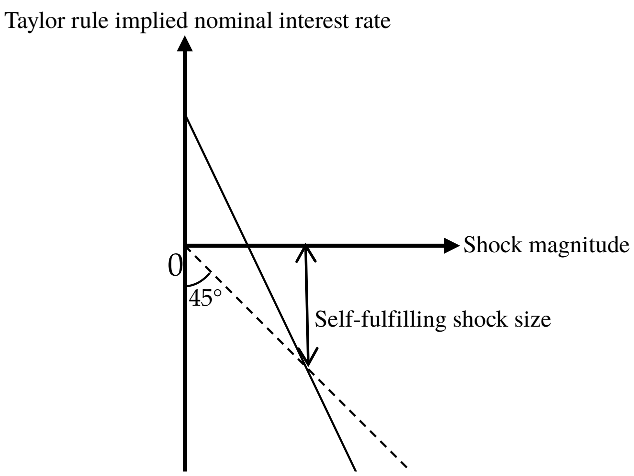


Figure 1: Self-fulfilling news shocks

Thinking in terms of endogenous news shocks also provides intuition for the presence of multiple equilibria in these models. As an example, consider a New Keynesian model with significant real and nominal frictions. If these frictions are large enough, then

[3] This has some similarities to the results of Richter and Throckmorton (2014) and Appendix B of Gavin et al. (2015), who show numerically that a particular solution algorithm does not converge in certain areas of the state/parameter/guess space for a simple NK model. However, our results are theoretical, so whereas Richter and Throckmorton and Gavin et al.'s results may possibly be driven by the particular properties of their solution procedure, ours imply true non-existence, at least for perfect foresight, otherwise linear models. For example, for the model with Rotemberg (1982) type pricing, and no steady-state distortions, that these authors work with, our results imply global existence and uniqueness for the linearized model when the standard Taylor principle is satisfied.
[4] The idea of imposing the zero lower bound by adding news shocks is also present in Holden (2010), Hebden et al. (2011), Holden and Paetz (2012) and Bodenstein et al. (2013). News shocks were introduced to the literature by Beaudry and Portier (2006).

learning about a future positive shock to nominal interest rates induces a sufficiently severe downtown that the Taylor rule calls for much lower rates, even in the period in which the shock actually arrives. While positive shocks having negative effects may sound somewhat bizarre, in fact this is a relatively common phenomenon in New Keynesian models. Then, there will be some magnitude of news shock to nominal interest rates today at which the news is of precisely the correct magnitude to bring the negative interest rates implied by the Taylor rule up to zero, in that period. A news shock of this magnitude thus becomes a self-fulfilling prophecy, as illustrated in Figure 1. In models with weaker rigidities, multiple equilibria are still possible if there is some combination of future periods such that with appropriate news shocks in each, a similar self-fulfilling prophecy occurs.

Our paper is structured as follows. In the following section, section 2 we present our key theoretical results on otherwise linear perfect foresight models. We then discuss the application of these results to New Keynesian models in section 3. Section 4 concludes. All files needed for the replication of this paper's numerical results are included in the "Examples" directory of the author's DynareOBC toolkit,[5] which implements an algorithm for simulating models with occasionally binding constraints that we discuss in a companion paper (Holden 2016).

2. Theoretical results on occasionally binding constraints in otherwise linear models under perfect foresight

In this section, we present our main theoretical results on existence and uniqueness of perfect foresight solutions to models which are linear apart from an occasionally binding constraint. We start by defining the problem to be solved, and examining its relationship both to the problem without OBCs, and to a related problem with news shocks to the bounded variable. Using the news shock representation, we demonstrate that solving the model with OBCs is equivalent to solving a linear complementarity problem. We then discuss some theoretical background on these problems, before presenting the main existence and uniqueness results.

2.1. Problem set-ups

We start by describing the problem set-up without bounds. Suppose that for $t \in \mathbb{N}^+$, (i.e. $t \in \mathbb{N}$, $t > 0$), the first order conditions of some model may be represented as:

$$\left(\hat{A} + \hat{B} + \hat{C}\right)\hat{\mu} = \hat{A}\hat{x}_{t-1} + B\hat{x}_t + \hat{C}\mathbb{E}_t\hat{x}_{t+1} + \hat{D}\varepsilon_t,$$

where $\hat{\mu} \in \mathbb{R}^{\hat{n}}$ and $\hat{x}_t \in \mathbb{R}^{\hat{n}}$, $\varepsilon_t \in \mathbb{R}^m$, $\mathbb{E}_{t-1}\varepsilon_t = 0$ for all $t \in \mathbb{N}^+$, and suppose that $\hat{x}_0$ is given as an initial condition. Throughout this paper, we will refer to first order

[5] These files may be viewed online at https://github.com/tholden/dynareOBC/tree/master/Examples.

conditions such as these as “the model”, conflating them with the optimisation problem(s) which gave rise to them.

Furthermore, suppose that $\varepsilon_t = 0$ for $t > 1$, as in an impulse response or perfect foresight simulation exercise. Additionally, we assume the existence of a terminal condition of the form $\hat{x}_t \to \hat{\mu}$ as $t \to \infty$, coming, for example, from the source model’s transversality constraints.

For $t \in \mathbb{N}^+$, define $x_t := \begin{bmatrix} \hat{x}_t \\ \varepsilon_{t+1} \end{bmatrix}$, $\mu := \begin{bmatrix} \hat{\mu} \\ 0 \end{bmatrix}$, $A := \begin{bmatrix} \hat{A} & \hat{D} \\ 0 & 0 \end{bmatrix}$, $B := \begin{bmatrix} \hat{B} & 0 \\ 0 & I \end{bmatrix}$, $C := \begin{bmatrix} \hat{C} & 0 \\ 0 & 0 \end{bmatrix}$, then, for $t \in \mathbb{N}^+$:

$$(A + B + C)\mu = Ax_{t-1} + Bx_t + Cx_{t+1}, \tag{1}$$

and we have the extended initial condition $x_0 = \begin{bmatrix} \hat{x}_0 \\ \varepsilon_1 \end{bmatrix}$, and the extended terminal condition $x_t \to \mu$ as $t \to \infty$. Expectations have disappeared since there is no uncertainty after period 0. Thus, the problem of solving the original model has the same form as that given in:

Problem 1 Suppose that $x_0 \in \mathbb{R}^n$ is given. Find $x_t \in \mathbb{R}^n$ for $t \in \mathbb{N}^+$ such that $x_t \to \mu$ as $t \to \infty$, and such that for all $t \in \mathbb{N}^+$, equation (1) holds.

We make the following assumption in all of the following:

Assumption 1 For any given $x_0 \in \mathbb{R}^n$, Problem 1 has a unique solution, which takes the form $x_t = (I - F)\mu + Fx_{t-1}$, for $t \in \mathbb{N}^+$, where $F = -(B + CF)^{-1}A$, and where all of the eigenvalues of F are weakly inside the unit circle.

Sims’s (2002) generalisation of the standard Blanchard-Kahn (1980) conditions is necessary and sufficient for this. Further, to avoid dealing specially with the knife-edge case of exact unit eigenvalues (even if they are constrained to the part of the model that is solved forward), in the following we rule it out with the subsequent assumption, which is, in any case, a necessary condition for perturbation to produce a consistent approximation to a source non-linear model, and which is also necessary for the linear model to have a unique steady-state:

Assumption 2 $\det(A + B + C) \neq 0$.

The combination of Assumption 1 and Assumption 2 imply that all of the eigenvalues of F are strictly inside the unit circle.

We are interested in models featuring occasionally binding constraints. We will concentrate on models featuring a single zero lower bound type constraint in their first equation, which we treat as defining the first element of x_t. Generalising from this special case to models with one or more fully general bounds is straightforward, and is discussed in the companion paper (Holden 2016). All of the results below go through in the more general case with minimal effort.

First, let us write $x_{1,t}$, $I_{1,\cdot}$, $A_{1,\cdot}$, $B_{1,\cdot}$, $C_{1,\cdot}$ for the first row of x_t, I, A, B, C (respectively) and $x_{-1,t}$, $I_{-1,\cdot}$, $A_{-1,\cdot}$, $B_{-1,\cdot}$, $C_{-1,\cdot}$ for the remainders. Likewise, we write $I_{\cdot,1}$ for the first column of I, and so on. Then, we are interested in the solution to:

Problem 2 Suppose that $x_0 \in \mathbb{R}^n$ is given. Find $T \in \mathbb{N}$ and $x_t \in \mathbb{R}^n$ for $t \in \mathbb{N}^+$ such that $x_t \to \mu$ as $t \to \infty$, and such that for all $t \in \mathbb{N}^+$:

$$x_{1,t} = \max\{0, I_{1,\cdot}\mu + A_{1,\cdot}(x_{t-1} - \mu) + (B_{1,\cdot} + I_{1,\cdot})(x_t - \mu) + C_{1,\cdot}(x_{t+1} - \mu)\},$$
$$(A_{-1,\cdot} + B_{-1,\cdot} + C_{-1,\cdot})\mu = A_{-1,\cdot}x_{t-1} + B_{-1,\cdot}x_t + C_{-1,\cdot}x_{t+1},$$

and such that $x_{1,t} > 0$ for $t > T$.

Note that in this problem we are implicitly ruling out any solutions which get permanently stuck at an alternative steady-state, by assuming that the terminal condition remains as before. In the monetary policy context, this amounts to assuming that the central banks' inflation target is credible. Since $x_{1,t} \to \mu_1$ as $t \to \infty$, it is without loss of generality to assume the existence of a $T \in \mathbb{N}$ such that $x_{1,t} > 0$ for $t > T$, but this T will play an important role in the below, so we introduce it now. We continue to assume that there is no uncertainty after period 0, so, in this non-linear model, the path of the endogenous variables will not necessarily match up with the path of their expectation in a richer model in which there was uncertainty after period 0.

In many models, the occasionally binding constraint comes from the KKT conditions of an optimisation problem. We will give in section 2.6 a general procedure for converting such conditions into a problem in the form of that Problem 2. The intuition is that one can use the model's equations to find the value the (lower) constrained variable would take were there no constraint and were the Lagrange multiplier on the constraint equal to zero today. This gives a "shadow" value of the constrained variable, and the actual value it takes will be the maximum of the bound and this shadow value.

We will analyse Problem 2 with the help of solutions to the auxiliary problem:

Problem 3 Suppose that $T \in \mathbb{N}$, $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^T$ is given. Find $x_t \in \mathbb{R}^n$, $y_t \in \mathbb{R}^T$ for $t \in \mathbb{N}^+$ such that $x_t \to \mu$, $y_t \to 0$, as $t \to \infty$, and such that for all $t \in \mathbb{N}^+$:

$$(A + B + C)\mu = Ax_{t-1} + Bx_t + Cx_{t+1} + I_{\cdot,1}y_{1,t-1},$$
$$y_{T,t} = 0, \qquad \forall i \in \{1, \dots, T-1\}, \qquad y_{i,t} = y_{i+1,t-1}.$$

This may be thought of as a version of Problem 1 with news shocks up to horizon T added to the first equation. The value of $y_{t,0}$ gives the news shock that hits in period t, i.e. $y_{1,t-1} = y_{t,0}$ for $t \le T$, and $y_{1,t-1} = 0$ for $t > T$.

2.2. Relationships between the problems

Since $y_{1,t-1} = 0$ for $t > T$, and using Assumption 1, $(x_{T+1} - \mu) = F(x_T - \mu)$, so with $t = T$, defining $s_{T+1} \coloneqq 0$, $(x_{t+1} - \mu) = s_{t+1} + F(x_t - \mu)$. Proceeding now by backwards induction on t, note that $0 = A(x_{t-1} - \mu) + B(x_t - \mu) + CF(x_t - \mu) + Cs_{t+1} + I_{\cdot,1}y_{t,0}$, so:

$$\begin{aligned}(x_t - \mu) &= -(B + CF)^{-1}\left[A(x_{t-1} - \mu) + Cs_{t+1} + I_{\cdot,1}y_{t,0}\right] \\ &= F(x_{t-1} - \mu) - (B + CF)^{-1}\left(Cs_{t+1} + I_{\cdot,1}y_{t,0}\right),\end{aligned}$$

i.e., if we define: $s_t \coloneqq -(B + CF)^{-1}\left(Cs_{t+1} + I_{\cdot,1}y_{t,0}\right)$, then $(x_t - \mu) = s_t + F(x_{t-1} - \mu)$. By induction then, this holds for all $t \in \{1, \dots, T\}$.[6] Hence, we have proved the following lemma:

Lemma 1 There is a unique solution to Problem 3 that is linear in x_0 and y_0.

For future reference, let $x_t^{(3,k)}$ be the solution to Problem 3 when $x_0 = \mu$, $y_0 = I_{\cdot,k}$ (i.e. a vector which is all zeros apart from a 1 in position k). Then, by linearity, for arbitrary y_0 the solution to Problem 3 when $x_0 = \mu$ is given by:

$$x_t - \mu = \sum_{k=1}^{T} y_{k,0}\left(x_t^{(3,k)} - \mu\right).$$

Let $M \in \mathbb{R}^{T\times T}$ satisfy:

$$M_{t,k} = x_{1,t}^{(3,k)} - \mu_1, \qquad \forall t,k \in \{1, .., T\}, \tag{2}$$

i.e. M horizontally stacks the (column-vector) relative impulse responses to the news shocks. Then this result implies that for arbitrary y_0, the path of the first variable in the solution to Problem 3 when $x_0 = \mu$ is given by: $(x_{1,1:T})' = \mu_1 + My_0$, where $x_{1,1:T}$ is the row vector of the first T values of the first component of x_t. Furthermore, for both arbitrary x_0 and y_0, the path of the first variable in the solution to Problem 3 is given by: $(x_{1,1:T})' = q + My_0$, where $q \coloneqq \left(x_{1,1:T}^{(1)}\right)'$ and $x_t^{(1)}$ is the unique solution to Problem 1, for the given x_0.[7]

Now let $x_t^{(2)}$ be a solution to Problem 2 given an arbitrary x_0. Since $x_t^{(2)} \to \mu$ as $t \to \infty$, there exists $T' \in \mathbb{N}$ such that for all $t > T'$, $x_{1,t}^{(2)} > 0$. We assume without loss of generality that $T' \leq T$. We seek to relate the solution to Problem 2 with the solution to Problem 3 for an appropriate choice of y_0. First, for all $t \in \mathbb{N}^+$, let:

$$e_t \coloneqq \begin{cases} -\left[I_{1,\cdot}\mu + A_{1,\cdot}\left(x_{t-1}^{(2)} - \mu\right) + \left(B_{1,\cdot} + I_{1,\cdot}\right)\left(x_t^{(2)} - \mu\right) + C_{1,\cdot}\left(x_{t+1}^{(2)} - \mu\right)\right] & \text{if } x_{1,t}^{(2)} = 0 \\ 0 & \text{if } x_{1,t}^{(2)} > 0 \end{cases}, \tag{3}$$

i.e. e_t is the shock that would need to hit the first equation for the positivity constraint on $x_{1,t}^{(2)}$ to be enforced. Note for future reference that by the definition of Problem 2,

[6] This representation of the solution to Problem 3 was inspired by that of Anderson (2015).

[7] This representation was also exploited by Holden (2010) and Holden and Paetz (2012).

$e_t \geq 0$ and $x_{1,t}^{(2)} e_t = 0$, for all $t \in \mathbb{N}^+$. From this definition, we also have that for all $t \in \mathbb{N}^+$, $0 = A\left(x_{t-1}^{(2)} - \mu\right) + B\left(x_t^{(2)} - \mu\right) + C\left(x_{t+1}^{(2)} - \mu\right) + I_{\cdot,1} e_t$. Furthermore, if $t > T$, then $t > T'$, and hence $e_t = 0$. Hence, by Assumption 1, $\left(x_{T+1}^{(2)} - \mu\right) = F\left(x_T^{(2)} - \mu\right)$. Thus, much as before, with $t = T$, defining $\tilde{s}_{T+1} \coloneqq 0$, $\left(x_{t+1}^{(2)} - \mu\right) = \tilde{s}_{t+1} + F\left(x_t^{(2)} - \mu\right)$. Consequently, $0 = A\left(x_{t-1}^{(2)} - \mu\right) + B\left(x_t^{(2)} - \mu\right) + CF\left(x_t^{(2)} - \mu\right) + C\tilde{s}_{t+1} + I_{\cdot,1} e_t$, so $\left(x_t^{(2)} - \mu\right) = F\left(x_{t-1}^{(2)} - \mu\right) - (B + CF)^{-1}(C\tilde{s}_{t+1} + I_{\cdot,1} e_t)$, i.e., if we define: $\tilde{s}_t \coloneqq -(B + CF)^{-1}(C\tilde{s}_{t+1} + I_{\cdot,1} e_t)$, then $\left(x_t^{(2)} - \mu\right) = \tilde{s}_t + F\left(x_{t-1}^{(2)} - \mu\right)$. As before, by induction this must hold for all $t \in \{1, \dots, T\}$. By comparing the definitions of s_t and $\tilde{s}_t$, and the laws of motion of x_t under both problems, we then immediately have that if Problem 3 is started with $x_0 = x_0^{(2)}$ and $y_0 = e_{1:T}'$, then $x_t^{(2)}$ solves Problem 3. Conversely, if $x_t^{(2)}$ solves Problem 3 for some y_0, then from the laws of motion of x_t under both problems it must be the case that $\tilde{s}_t = s_t$ for all $t \in \mathbb{N}$, and hence from the definitions of s_t and $\tilde{s}_t$, we have that $y_0 = e_{1:T}'$. This has established the following result:

Lemma 2 For any solution, $x_t^{(2)}$ to Problem 2:

1) With $e_{1:T}$ as defined in equation (3), $e_{1:T} \geq 0$, $x_{1,1:T}^{(2)} \geq 0$ and $x_{1,1:T}^{(2)} \circ e_{1:T} = 0$, where $\circ$ denotes the Hadamard (entry-wise) product.
2) $x_t^{(2)}$ is also the unique solution to Problem 3 with $x_0 = x_0^{(2)}$ and $y_0 = e_{1:T}'$.
3) If $x_t^{(2)}$ solves Problem 3 with $x_0 = x_0^{(2)}$ and with some y_0, then $y_0 = e_{1:T}'$.

However, to use the easy solution to Problem 3 to assist us in solving Problem 2 requires a slightly stronger result. Suppose that $y_0 \in \mathbb{R}^T$ is such that $y_0 \geq 0$, $x_{1,1:T}^{(3)} \circ y_0' = 0$ and $x_{1,t}^{(3)} \geq 0$ for all $t \in \mathbb{N}$, where $x_t^{(3)}$ is the unique solution to Problem 3 when started at x_0, y_0. We would like to prove that in this case $x_t^{(3)}$ must also be a solution to Problem 2. I.e., we must prove that for all $t \in \mathbb{N}^+$:

$$x_{1,t}^{(3)} = \max\left\{0, I_{1,\cdot}\mu + A_{1,\cdot}\left(x_{t-1}^{(3)} - \mu\right) + (B_{1,\cdot} + I_{1,\cdot})\left(x_t^{(3)} - \mu\right) + C_{1,\cdot}\left(x_{t+1}^{(3)} - \mu\right)\right\}, \quad (4)$$

$$(A_{-1,\cdot} + B_{-1,\cdot} + C_{-1,\cdot})\mu = A_{-1,\cdot} x_{t-1}^{(3)} + B_{-1,\cdot} x_t^{(3)} + C_{-1,\cdot} x_{t+1}^{(3)}.$$

By the definition of Problem 3, the latter equation must hold with equality, so there is nothing to prove there. Hence we just need to prove that equation (4) holds for all $t \in \mathbb{N}^+$. So let $t \in \mathbb{N}^+$. Now, if $x_{1,t}^{(3)} > 0$, then $y_{t,0} = 0$, by the complementary slackness type condition ($x_{1,1:T}^{(3)} \circ y_0' = 0$). Thus, from the definition of Problem 3:

$$\begin{aligned} x_{1,t}^{(3)} &= I_{1,\cdot}\mu + A_{1,\cdot}\left(x_{t-1}^{(3)} - \mu\right) + (B_{1,\cdot} + I_{1,\cdot})\left(x_t^{(3)} - \mu\right) + C_{1,\cdot}\left(x_{t+1}^{(3)} - \mu\right) \\ &= \max\left\{0, I_{1,\cdot}\mu + A_{1,\cdot}\left(x_{t-1}^{(3)} - \mu\right) + (B_{1,\cdot} + I_{1,\cdot})\left(x_t^{(3)} - \mu\right) + C_{1,\cdot}\left(x_{t+1}^{(3)} - \mu\right)\right\}, \end{aligned}$$

as required. The only remaining case is that $x_{1,t}^{(3)} = 0$ (since $x_{1,t}^{(3)} \geq 0$ for all $t \in \mathbb{N}$, by assumption), which implies that:

$$\begin{aligned} x_{1,t}^{(3)} = 0 &= A_{1,\cdot}(x_{t-1} - \mu) + B_{1,\cdot}(x_t - \mu) + C_{1,\cdot}(x_{t+1} - \mu) + y_{t,0} \\ &= I_{1,\cdot}\mu + A_{1,\cdot}(x_{t-1} - \mu) + (B_{1,\cdot} + I_{1,\cdot})(x_t - \mu) + C_{1,\cdot}(x_{t+1} - \mu) + y_{t,0}, \end{aligned}$$

by the definition of Problem 3. Thus:

$$I_{1,\cdot}\mu + A_{1,\cdot}(x_{t-1}-\mu) + (B_{1,\cdot}+I_{1,\cdot})(x_t-\mu) + C_{1,\cdot}(x_{t+1}-\mu) = -y_{t,0} \le 0,$$

where the inequality is an immediate consequence of another of our assumptions. Consequently, equation (4) holds in this case too. Together with Lemma 1, Lemma 2, and our representation of the solution of Problem 3, this completes the proof of the following proposition:

Proposition 1 The following hold:

1) Let $x_t^{(3)}$ be the unique solution to Problem 3 when initialized with some x_0, y_0. Then $x_t^{(3)}$ is a solution to Problem 2 when initialized with x_0 if and only if $y_0 \ge 0$, $y_0 \circ (q + My_0) = 0$, $q + My_0 \ge 0$ and $x_{1,t}^{(3)} \ge 0$ for all $t \in \mathbb{N}$ with $t > T$.
2) Let $x_t^{(2)}$ be any solution to Problem 2 when initialized with x_0. Then there exists a $y_0 \in \mathbb{R}^T$ such that $y_0 \ge 0$, $y_0 \circ (q + My_0) = 0$, $q + My_0 \ge 0$, such that $x_t^{(2)}$ is the unique solution to Problem 3 when initialized with x_0, y_0.

2.3. The linear complementarity representation

Proposition 1 establishes that providing we initially choose T sufficiently high, to find a solution to Problem 2, it is sufficient to solve the following problem instead:

Problem 4 Suppose $q \in \mathbb{R}^T$ and $M \in \mathbb{R}^{T\times T}$ are given. Find $y \in \mathbb{R}^T$ such that $y \ge 0$, $y \circ (q + My) = 0$ and $q + My \ge 0$. We call this the **linear complementarity problem (LCP)** (q, M). (Cottle 2009)

These problems have been extensively studied, and so we can import results on the properties of LCPs to derive results on the properties of solutions to models with OBCs.

All of the results in the mathematical literature rest on properties of the matrix M, thus we would like to establish if the structure of our particular M implies it has any special properties. Unfortunately, we prove the following result in this paper's companion paper (Holden 2016), which implies that M has no general properties:

Proposition 2 For any matrix $\mathscr{M} \in \mathbb{R}^{T\times T}$, there exists a model in the form of Problem 2 with a number of state variables given by a quadratic in T, such that $M = \mathscr{M}$ for that model, where M is defined as in equation (2), and such that for all $\mathscr{q} \in \mathbb{R}^T$, there exists an initial state for which $q = \mathscr{q}$, where q is the path of the bounded variable when constraints are ignored. (Holden 2016)

We now introduce some definitions of matrix properties that are necessary for the statement of our key existence and uniqueness results. The ultimate properties of the solutions to the OBC model are determined by which of these matrix properties M possesses. In each case, we give the definitions in a constructive form which makes clear both how the property might be verified computationally, and the links between definitions. These are not necessarily in the form which is standard in the original literature, however. For both the original definitions, and the proofs of equivalence between the ones below and the originals, see Cottle, Pang, and Stone (2009a) and Xu (1993) (for the characterisation of sufficient models).

Definition 1 (Principal sub-matrix, Principal minor) For a matrix $M \in \mathbb{R}^{T\times T}$, the **principal sub-matrices** of M are the matrices:

$$\left\{\left[M_{i,j}\right]_{i,j=k_1,\dots,k_S}\middle|S, k_1, \dots, k_S \in \{1,\dots,T\}, k_1 < k_2 < \cdots < k_S\right\},$$

i.e. the **principal sub-matrices** of M are formed by deleting the same rows and columns. The **principal minors** of M are the collection of values:

$$\left\{\det\left(\left[M_{i,j}\right]_{i,j=k_1,\dots,k_S}\right)\middle|S, k_1, \dots, k_S \in \{1,\dots,T\}, k_1 < k_2 < \cdots < k_S\right\},$$

i.e. the **principal minors** of M are the determinants of the principal sub-matrices of M.

Definition 2 (P($_0$)-matrix) A matrix $M \in \mathbb{R}^{T\times T}$ is called a **P-matrix (P_0-matrix)** if the principal minors of M are all strictly (weakly) positive. *Note: for symmetric M, M is a P($_0$)-matrix if and only if all of its eigenvalues are strictly (weakly) positive.*

Definition 3 (General positive (semi-)definite) A matrix $M \in \mathbb{R}^{T\times T}$ is called **general positive (semi-)definite** if $M + M'$ is a P-matrix (P_0-matrix). *If M is symmetric, then, M is general positive (semi-)definite if and only if it is positive (semi-)definite.*

Definition 4 (Sufficient matrices) Let $M \in \mathbb{R}^{T\times T}$. M is called **column sufficient** if M is a P_0-matrix, and for each principal sub-matrix $W \coloneqq \left[M_{i,j}\right]_{i,j=k_1,\dots,k_S}$ of M, with zero determinant, and for each proper principal sub-matrix $\left[W_{i,j}\right]_{i,j=l_1,\dots,l_R}$ of W ($R < S$), with zero determinant, the columns of $\left[W_{i,j}\right]_{\substack{i=1,\dots,S\\ j=l_1,\dots,l_R}}$ do not form a basis for the column space of W.[8] M is called **row sufficient** if M' is column sufficient. M is called **sufficient** if it is column sufficient and row sufficient.

Definition 5 (S($_0$)-matrix) A matrix $M \in \mathbb{R}^{T\times T}$ is called an **S-matrix (S_0-matrix)** if there exists $y \in \mathbb{R}^T$ such that $y > 0$ and $My \gg 0$ ($My \geq 0$).[9]

Definition 6 ((Strictly) Semi-monotone) A matrix $M \in \mathbb{R}^{T\times T}$ is called **(strictly) semi-monotone** if each of its principal sub-matrices is an **S_0-matrix (S-matrix)**.

---

[8] This may be checked via the singular value decomposition.

[9] These condition may be rewritten as $\sup\{\varsigma \in \mathbb{R} | \exists y \geq 0 \text{ s.t. } \forall t \in \{1,\dots,T\}, (My)_t \geq \varsigma \wedge y_t \leq 1\} > 0$, and $\sup\left\{\sum_{t=1}^T y_t \middle| y \geq 0, My \geq 0 \wedge \forall t \in \{1,\dots,T\}, y_t \leq 1\right\} > 0$, respectively. As linear-programming problems, these may be verified in time polynomial in T using the methods described in e.g. Roos, Terlaky, and Vial (2006). Alternatively, by Ville's theorem of the alternative (Cottle, Pang, and Stone 2009b), M is not an S_0-matrix if and only if $-M'$ is an S-matrix.

Definition 7 ((Strictly) Copositive) A matrix $M \in \mathbb{R}^{T \times T}$ is called (**strictly**) **copositive** if $M + M'$ is (strictly) semi-monotone.[10]

Cottle, Pang, and Stone (2009a) note the following relationships between these classes (amongst others):

Lemma 3 The following hold:

1) All general positive semi-definite matrices are copositive and sufficient.
2) P_0 includes skew-symmetric matrices, general positive semi-definite matrices, sufficient matrices and P-matrices.
3) All P_0-matrices, and all copositive matrices are semi-monotone, and all P-matrices, and all strictly copositive matrices are strictly semi-monotone (and hence also S-matrices).

Additionally, from considering the 1×1 principal sub-matrices of M, we have the following restrictions on the diagonal of M:

Lemma 4 All general positive semi-definite, semi-monotone, sufficient, P_0 and copositive matrices have non-negative diagonals, and all general positive definite, strictly semi-monotone, P and strictly copositive matrices have positive diagonals.

For many macroeconomic models, this simple condition is sufficient to rule out membership of these matrix classes, as medium-scale DSGE models[11] with a ZLB frequently have negative elements on the diagonal of their M matrix, when T is large enough. Thus, following the intuition of Figure 1, such models will satisfy the conditions to have multiple equilibria, though they will not be the only such models.

Unfortunately, for all of these matrix classes except the classes of general positive (semi-)definite matrices, and $S_{(0)}$-matrices, no algorithm which runs in an amount of time that is polynomial in T is known, thus verifying class membership may not be feasible with large T. However, disproving class membership only requires finding one principal sub-matrix which fails to have the required property, and for this, starting with the 1×1 principal sub-matrices (e.g. the diagonal), then considering the 2×2 ones (etc.) is often a good strategy.[12]

A common intuition is that in models without state variables, M must be both a P matrix, and an S matrix. In fact, this is not true. Indeed, there are even purely static

[10] Väliaho (1986) contains an alternative characterisation which avoids solving any linear programming problems.

[11] This applies, for example, to the Smets and Wouters (2003) model, as we will show in section 3.5.

[12] The facts that all of the eigenvalues of a $T \times T$ P-matrix have complex arguments in the interval $(-\pi + \frac{\pi}{T}, \pi - \frac{\pi}{T})$, and all of the eigenvalues of a $T \times T$ P_0-matrix have complex arguments in the interval $[-\pi + \frac{\pi}{T}, \pi - \frac{\pi}{T}]$ (Fang 1989) may also assist in ruling out these matrix classes.

models for which M is not in either of these classes. For example, in online appendix A, we construct a purely static model for which $M_{1:\infty,1:\infty} = -I_{\infty\times\infty}$, which is neither a P-matrix, nor an S-matrix, for any T.

2.4. Existence results

We start by considering necessary or sufficient conditions for the existence of a solution to a model with occasionally binding constraints. Ideally, we would like the solution to exist for any possible path the bounded variable might have taken in the future were there no OBC, i.e. for any possible q. To see this, note that under a perfect foresight exercise we are ignoring the fact that shocks might hit the economy in future. More properly, we ought to integrate over future uncertainty, as in the stochastic extended path approach of Adjemian and Juillard (2013). A crude way to do this would just be to draw lots of samples of future shocks for periods $1, \dots, S$, and average over these draws. However, in a linear model with shocks with unbounded support, providing at least one shock has an impact on a given variable, the distribution of future paths of that variable has positive support over the entirety of $\mathbb{R}^S$. Thus, ideally we would like M to be such that for any q, the linear complementarity problem (q, M) has a solution.

Definition 8 (Feasible LCP) Suppose $q \in \mathbb{R}^T$ and $M \in \mathbb{R}^{T\times T}$ are given. The LCP corresponding to M and q is called **feasible** if there exists $y \in \mathbb{R}^T$ such that $y \geq 0$ and $q + My \geq 0$.

By construction, if an LCP (q, M) has a solution, then it is feasible, i.e. being feasible is a necessary condition for existence. Checking feasibility is straightforward for any particular (q, M), since to find a feasible solution we just need to solve a standard linear programming problem, which is possible in an amount of time that is polynomial in T.

Note that if the LCP (q, M) is not feasible, then for any $\hat{q} \leq q$, if $y \geq 0$, then $\hat{q} + My \leq q + My < 0$ since (q, M) is not feasible, so the LCP $(\hat{q}, M)$ is also not feasible. Consequently, if there are any q for which the LCP is non-feasible, then there is a positive measure of such q. Thus, in a model with uncertainty, if there are some q for which the model has no solution satisfying the terminal condition, even with arbitrarily large T, then the model will have no solution satisfying the terminal condition with positive probability. This in turn means that it is not consistent with rationality for agents to believe that our terminal condition is satisfied with certainty, so they would have to place some positive probability on getting stuck in an alternative steady-state.

The following proposition gives an easily verified necessary condition for the global existence of a solution to the model with occasionally binding constraints, given some fixed horizon T:

Proposition 3 The LCP (q, M) is feasible for all $q \in \mathbb{R}^T$ if and only if M is an S-matrix. (Cottle, Pang, and Stone 2009a) [13]

Of course, it may be the case that the M matrix is only an S-matrix when T is very large, so we must be careful in using this condition to imply non-existence of a solution. Furthermore, it may be the case that although there exists some $y \in \mathbb{R}^T$ with $y \geq 0$ such that $M_{1:T,1:T}y \gg 0$, where we are indexing the M matrix by its size for clarity, for any such y, $\inf_{t \in \mathbb{N}^+} M_{t,1:T}y < 0$, so for some $q \in \mathbb{R}^{\mathbb{N}^+}$, the infinite LCP $(q, M_{1:\infty,1:\infty})$ is not feasible under the additional restriction that $y_t = 0$ for $t > T$. Strictly, it is this infinite LCP which we ought to be solving, subject to the additional constraint that y has only finitely many non-zero elements, as implied by our terminal condition.

By Proposition 3, this infinite problem is feasible if and only if:

$$\varsigma \coloneqq \sup_{\substack{y \in [0,1]^{\mathbb{N}^+} \\ \exists T \in \mathbb{N} \text{ s.t. } \forall t > T, y_t = 0}} \inf_{t \in \mathbb{N}^+} M_{t,1:\infty}y > 0.$$

Consequently, if $\varsigma > 0$ then for every $q \in \mathbb{R}^{\mathbb{N}^+}$, for sufficiently large T, the finite problem $(q_{1:T}, M_{1:T,1:T})$ will be feasible, which is a sufficient condition for solvability. In order to evaluate this limit, we first need to derive constructive bounds on the M matrix for large T. We do this in the online appendix B, where we prove that the rows and columns of M are converging to 0 (with constructive bounds), and that the k^{th} diagonal of the M matrix is converging to the value $d_{1,k}$, to be defined (again with constructive bounds), where diagonals are indexed such that the principal diagonal is index 0, and indices increase as one moves up and to the right in the M matrix. To explain the origins of $d_{1,k}$ we note the following lemma proved in online appendix B:

Lemma 5 The (time-reversed) difference equation $A\hat{d}_{k+1} + B\hat{d}_k + C\hat{d}_{k-1} = 0$ for all $k \in \mathbb{N}^+$ has a unique solution satisfying the terminal condition $\hat{d}_k \to 0$ as $k \to \infty$, given by $\hat{d}_k = H\hat{d}_{k-1}$, for all $k \in \mathbb{N}^+$, for some H with eigenvalues in the unit circle.

Then, we define $d_0 \coloneqq -(AH + B + CF)^{-1}I_{\cdot,1}$, $d_k = Hd_{k-1}$, for all $k \in \mathbb{N}^+$, and $d_{-t} = Fd_{-(t-1)}$, for all $t \in \mathbb{N}^+$, so d_k follows the time reversed difference equation for positive indices, and the original difference equation for negative indices. This is opposite to

[13] Most of the results on LCPs in both this and the following section are restatements of (assorted) results contained in Cottle, Pang, and Stone (2009a) and Väliaho (1986) (for the characterisation of "copositive-plus" matrices), and the reader is referred to those works for proofs and further references.

what one might perhaps expect since time is increasing as one descends the rows of M, but diagonal indices are decreasing as one descends in M.

Using the resulting bounds on M, we can construct upper and lower bounds on ς, which are described in the following propositions, also proven in online appendix B:

Proposition 4 There exists $\underline{\varsigma}_T, \overline{\varsigma}_T \geq 0$, defined in the online appendix B, computable in time polynomial in T, such that $\underline{\varsigma}_T \leq \varsigma \leq \overline{\varsigma}_T$, and $\left|\underline{\varsigma}_T - \overline{\varsigma}_T\right| \to 0$ as $T \to \infty$.

These conditions give simple tests for feasibility or non-feasibility with sufficiently large T.

We now turn to sufficient conditions for the existence of a solution for some finite T.

Proposition 5 The LCP (q, M) is solvable if it is feasible and, either:

1. M is row-sufficient, or,
2. M is copositive and for all non-singular principal sub-matrices W of M, all non-negative columns of W^{-1} possess a non-zero diagonal element.

(Cottle, Pang, and Stone 2009a; Väliaho 1986)

If either condition 1 or condition 2 of Proposition 5 is satisfied, then to check existence for any particular q, we only need to solve a linear programming problem to see if a solution exists for a particular q. As this may be substantially faster than solving the LCP, this may be helpful in practice.

Proposition 6 The LCP (q, M) is solvable for all $q \in \mathbb{R}^T$, if at least one of the following conditions holds:

1. M is an S-matrix, and either condition 1 or condition 2 of Proposition 5 are satisfied.
2. M is copositive with non-zero principal minors.
3. M is a P-matrix, a strictly copositive matrix or a strictly semi-monotone matrix.

(Cottle, Pang, and Stone 2009a)

If condition 1, 2 or 3 of Proposition 6 is satisfied, then we know that the LCP will always have a solution. Therefore, for any path of the bounded variable in the absence of the bound, we will also be able to solve the model when the bound is imposed. Monetary policy makers should always choose a policy rule that produces a model that satisfies one of these three conditions, if they can, since otherwise there is a positive probability that only solutions converging to the "bad" steady-state will exist in some state of the world.

Ideally, we might have liked conditions for the existence of a solution that are both necessary and sufficient, but unfortunately at present no such conditions exist in full generality. However, in the special case of M matrices with nonnegative entries, we have the following result:

Proposition 7 If M is a matrix with nonnegative entries, then the LCP (q, M) is solvable for all $q \in \mathbb{R}^T$, if and only if M has a strictly positive diagonal. (Cottle, Pang, and Stone 2009a)

2.5. Uniqueness results

While no fully general necessary and sufficient conditions have been derived for existence, such conditions are available for uniqueness, in particular:

Proposition 8 The LCP (q, M) has a unique solution for all $q \in \mathbb{R}^T$, if and only if M is a P-matrix. If M is not a P-matrix, then the LCP (q, M) has multiple solutions for some q. (Samelson, Thrall, and Wesler 1958; Cottle, Pang, and Stone 2009a)

This proposition is the equivalent for models with OBCs of the key proposition of Blanchard and Kahn (1980). By testing whether our matrix M is a P-matrix we can immediately determine if the model possesses a unique solution in any state of the world, and for any sequence of future shocks, for a fixed T. In our experience, this condition is satisfied in efficient models, such as models of irreversible investment, as one would expect, but is not generally satisfied in medium-scale New-Keynesian models with a ZLB on nominal interest rates. Given that if M is a P-matrix, so too are all its principal sub-matrices, if we see that M is not a P-matrix for some T, then we know that with larger T it would also not be a P-matrix. Thus, if for some T, M is not a P-matrix, then we know that the model does not have a unique solution, even for arbitrarily large T. Alternatively, we can prove that with large T some M is not a P-matrix by using the analytic formula for the limit of its diagonal given in the previous section, i.e. $d_{0,1} = -I_{1,\cdot}(AH + B + CF)^{-1} I_{\cdot,1}$. If this value is negative, then we know that with sufficiently large T, M will not be a P-matrix.

Since some classes of models almost never possess a unique solution when at the zero lower bound, we might reasonably require a lesser condition, namely that at least when the solution to the model without a bound is a solution to the model with the bound, then it ought to be the unique solution. This is equivalent to requiring that when q is non-negative, the LCP (q, M) has a unique solution. Conditions for this are given in the following proposition:

Proposition 9 The LCP (q, M) has a unique solution for all $q \in \mathbb{R}^T$ with $q \gg 0$ ($q \geq 0$) if and only if M is (strictly) semi-monotone. (Cottle, Pang, and Stone 2009a)

Hence, by verifying that M is (strictly) semi-monotone, we can reassure ourselves that merely introducing the bound will not change the solution away from the bound. When this condition is violated, even when the economy is a long way from the bound, there may be solutions which jump to the bound. Again, since principal sub-matrices of (strictly) semi-monotone are (strictly) semi-monotone, a failure of (strict) semi-monotonicity for some T implies a failure for all larger T. Furthermore, if $d_{0,1} < 0$ then again for sufficiently large T, M cannot be semi-monotone.

Where there are multiple solutions, we might like to be able to select one via some objective function. This is particularly tractable when either the number of solutions is finite, or the solution set is convex. Conditions for this are given in online appendix C.

2.6. Results from dynamic programming

Alternative existence and uniqueness results for the infinite T problem can be established via dynamic programming methods, under the assumption that Problem 2 comes from the first order conditions solution of a social planner problem. These have the advantage that their conditions are potentially much easier to evaluate, though they also have somewhat limited applicability. We focus here on uniqueness results, since these are generally of greater interest.

Suppose that the social planner in some economy solves the following problem:

Problem 5 Suppose $\mu \in \mathbb{R}^n$, $\Psi^{(0)} \in \mathbb{R}^{c\times 1}$ and $\Psi^{(1)} \in \mathbb{R}^{c\times 2n}$ are given, where $c \in \mathbb{N}$. Define $\widetilde{\Gamma}: \mathbb{R}^n \to \mathbb{P}(\mathbb{R}^n)$ (where $\mathbb{P}$ denotes the power-set operator) by:

$$\widetilde{\Gamma}(x) = \left\{ z \in \mathbb{R}^n \middle| \ 0 \leq \Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x-\mu \\ z-\mu \end{bmatrix} \right\}, \tag{5}$$

for all $x \in \mathbb{R}^n$. *(Note: $\widetilde{\Gamma}(x)$ will give the set of feasible values for next period's state if the current state is x. Equality constraints may be included by including an identical lower bound and upper bound.)* Define:

$$\tilde{X} := \left\{ x \in \mathbb{R}^n \middle| \widetilde{\Gamma}(x) \neq \emptyset \right\}, \tag{6}$$

and suppose without loss of generality that for all $x \in \mathbb{R}^n$, $\widetilde{\Gamma}(x) \cap \tilde{X} = \widetilde{\Gamma}(x)$. *(Note: this means that the linear inequalities bounding $\tilde{X}$ are already included in those in the definition of $\widetilde{\Gamma}(x)$. It is without loss of generality as the planner will never choose an $\tilde{x} \in \widetilde{\Gamma}(x)$ such that $\widetilde{\Gamma}(\tilde{x}) = \emptyset$.)* Further define $\widetilde{\mathcal{F}}: \tilde{X} \times \tilde{X} \to \mathbb{R}$ by:

$$\widetilde{\mathcal{F}}(x,z) = u^{(0)} + u^{(1)} \begin{bmatrix} x-\mu \\ z-\mu \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x-\mu \\ z-\mu \end{bmatrix}' \tilde{u}^{(2)} \begin{bmatrix} x-\mu \\ z-\mu \end{bmatrix}, \tag{7}$$

for all $x, z \in \tilde{X}$, where $u^{(0)} \in \mathbb{R}$, $u^{(1)} \in \mathbb{R}^{1\times 2n}$ and $\tilde{u}^{(2)} = \tilde{u}^{(2)\prime} \in \mathbb{R}^{2n\times 2n}$ are given. Finally, suppose $x_0 \in \tilde{X}$ is given and $\beta \in (0,1)$, and choose $x_1, x_2, \dots$ to maximise:

$$\liminf_{T\to\infty} \sum_{t=1}^{T} \beta^{t-1} \widetilde{\mathcal{F}}(x_{t-1}, x_t) \tag{8}$$

subject to the constraints that for all $t \in \mathbb{N}^+$, $x_t \in \widetilde{\Gamma}(x_{t-1})$.

To ensure the problem is well behaved, we make the following assumption:

Assumption 3 $\tilde{u}^{(2)}$ is negative-definite.

In online appendix D, we establish the following (unsurprising) result:

Proposition 10 If either $\tilde{X}$ is compact, or, $\widetilde{\Gamma}(x)$ is compact valued and $x \in \widetilde{\Gamma}(x)$ for all $x \in \tilde{X}$, then for all $x_0 \in \tilde{X}$, there is a unique path $(x_t)_{t=0}^{\infty}$ which solves Problem 5.

We wish to use this result to establish the uniqueness of the solution to the first order conditions. The Lagrangian for our problem is given by:

$$\sum_{t=1}^{\infty} \beta^{t-1} \left[\widetilde{\mathscr{F}}(x_{t-1}, x_t) + \lambda'_{\Psi,t} \left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} \right] \right], \tag{9}$$

for some KKT-multipliers $\lambda_t \in \mathbb{R}^c$ for all $t \in \mathbb{N}^+$. Taking the first order conditions leads to the following necessary KKT conditions, for all $t \in \mathbb{N}^+$:

$$0 = u^{(1)}_{\cdot,2} + \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}' \tilde{u}^{(2)}_{\cdot,2} + \lambda'_t \Psi^{(1)}_{\cdot,2} + \beta \left[u^{(1)}_{\cdot,1} + \begin{bmatrix} x_t - \mu \\ x_{t+1} - \mu \end{bmatrix}' \tilde{u}^{(2)}_{\cdot,1} + \lambda'_{t+1} \Psi^{(1)}_{\cdot,1} \right], \tag{10}$$

$$0 \le \Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}, \qquad 0 \le \lambda_t, \qquad 0 = \lambda_t \circ \left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} \right], \tag{11}$$

where subscripts 1 and 2 refer to blocks of rows or columns of length n. Additionally, for μ to be the steady-state of x_t and $\overline{\lambda}$ to be the steady-state of λ_t, we require:

$$0 = u^{(1)}_{\cdot,2} + \overline{\lambda}' \Psi^{(1)}_{\cdot,2} + \beta \left[u^{(1)}_{\cdot,1} + \overline{\lambda}' \Psi^{(1)}_{\cdot,1} \right], \tag{12}$$

$$0 \le \Psi^{(0)}, \qquad 0 \le \overline{\lambda}, \qquad 0 = \overline{\lambda} \circ \Psi^{(0)}. \tag{13}$$

In online appendix E we prove the following result:

Proposition 11 Suppose that for all $t \in \mathbb{N}$, $(x_t)_{t=1}^{\infty}$ and $(\lambda_t)_{t=1}^{\infty}$ satisfy the KKT conditions given in equations (10) and (11), and that as $t \to \infty$, $x_t \to \mu$ and $\lambda_t \to \overline{\lambda}$, where μ and λ satisfy the steady-state KKT conditions given in equations (12) and (13). Then $(x_t)_{t=1}^{\infty}$ solves Problem 5. If, further, either condition of Proposition 10 is satisfied, then $(x_t)_{t=1}^{\infty}$ is the unique solution to Problem 5, and there can be no other solutions to the KKT conditions given in equations (10) and (11) satisfying $x_t \to \mu$ and $\lambda_t \to \overline{\lambda}$ as $t \to \infty$.

Now, it is possible to convert the KKT conditions given in equations (10) and (11) into a problem in the form of the multiple-bound generalisation of Problem 2 quite generally. To see this, first note that we may rewrite equation (10) as:

$$\begin{aligned} 0 = u^{(1)\prime}_{\cdot,2} + \tilde{u}^{(2)}_{2,1}(x_{t-1} - \mu) + \tilde{u}^{(2)}_{2,2}(x_t - \mu) + \Psi^{(1)\prime}_{\cdot,2} \lambda_t \\ + \beta \left[u^{(1)\prime}_{\cdot,1} + \tilde{u}^{(2)}_{1,1}(x_t - \mu) + \tilde{u}^{(2)}_{1,2}(x_{t+1} - \mu) + \Psi^{(1)\prime}_{\cdot,1} \lambda_{t+1} \right]. \end{aligned}$$

Now, $\tilde{u}^{(2)}_{2,2} + \beta u^{(2)}_{1,1}$ is negative definite, hence it is valid to define:

$$\mathscr{V} := \Psi^{(1)}_{\cdot,2} \left[\tilde{u}^{(2)}_{2,2} + \beta \tilde{u}^{(2)}_{1,1} \right]^{-1}.$$

Then, equation (9) implies that:

$$\begin{aligned}
&\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} \\
&= \Psi^{(0)} + \left(\Psi_{\cdot,1}^{(1)} - \mathscr{V} \tilde{u}_{2,1}^{(2)}\right)(x_{t-1} - \mu) - \mathscr{V} \left[u_{\cdot,2}^{(1)\prime} + \beta \left[u_{\cdot,1}^{(1)\prime} + \tilde{u}_{1,2}^{(2)}(x_{t+1} - \mu) + \Psi_{\cdot,1}^{(1)\prime} \lambda_{t+1} \right]\right] \\
&-\Psi_{\cdot,2}^{(1)} \left[\tilde{u}_{2,2}^{(2)} + \beta \tilde{u}_{1,1}^{(2)}\right]^{-1} \Psi_{\cdot,2}^{(1)\prime} \lambda_t.
\end{aligned} \tag{14}$$

Moreover, equation (11) implies that if the k^{th} element of $\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}$ is strictly positive, then the k^{th} element of λ_t is zero, so:

$$\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} = \max\{0, z_t\}, \tag{15}$$

where:

$$\begin{aligned}
z_t \coloneqq \Psi^{(0)} &+ \left(\Psi_{\cdot,1}^{(1)} - \mathscr{V} \tilde{u}_{2,1}^{(2)}\right)(x_{t-1} - \mu) \\
&- \mathscr{V} \left[u_{\cdot,2}^{(1)\prime} + \beta \left[u_{\cdot,1}^{(1)\prime} + \tilde{u}_{1,2}^{(2)}(x_{t+1} - \mu) + \Psi_{\cdot,1}^{(1)\prime} \lambda_{t+1} \right]\right] \\
&- \left[\Psi_{\cdot,2}^{(1)} \left[\tilde{u}_{2,2}^{(2)} + \beta \tilde{u}_{1,1}^{(2)}\right]^{-1} \Psi_{\cdot,2}^{(1)\prime} + \mathscr{W} \right] \lambda_t,
\end{aligned}$$

and $\mathscr{W} \in \mathbb{R}^{c \times c}$ is an arbitrary, strictly positive diagonal matrix. A natural choice is:

$$\mathscr{W} \coloneqq -\operatorname{diag}\operatorname{diag}\left[\Psi_{\cdot,2}^{(1)} \left[\tilde{u}_{2,2}^{(2)} + \beta \tilde{u}_{1,1}^{(2)}\right]^{-1} \Psi_{\cdot,2}^{(1)\prime} \right],$$

providing this is strictly positive (it is weakly positive at least as $\tilde{u}_{2,2}^{(2)} + \beta \tilde{u}_{1,1}^{(2)}$ is negative definite), where the diag operator maps matrices to a vector containing their diagonal, and maps vectors to a matrix with the given vector on the diagonal, and zeros elsewhere.

We claim that we may replace equation (11) with equation (15) without changing the model. We have already shown that equation (11) implies equation (15), so we just have to prove the converse. We continue to suppose equation (9) holds, and thus, so too does equation (14). Then, from subtracting equation (14) from equation (15), we have that:

$$\mathscr{W} \lambda_t = \max\{-z_t, 0\}.$$

Hence, as $\mathscr{W}$ is a strictly positive diagonal matrix, and the right hand side is weakly positive, $\lambda_t \geq 0$. Furthermore, the kth element of λ_t is non-negative if and only if the kth element of z_t is non-positive (as $\mathscr{W}$ is a strictly positive diagonal matrix), which in turn holds if and only if the kth element of $\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}$ is equal to zero, by equation (15). Thus equation (11) is satisfied.

Combined with our previous results, this gives the following proposition:

Proposition 12 Suppose we are given a problem in the form of Problem 5. Then, the KKT conditions of that problem may be placed into the form of the multiple-bound generalisation of Problem 2. Let $\left(q_{x_0}, M\right)$ be the infinite LCP corresponding to this representation, given initial state $x_0 \in \tilde{X}$. Then, if y is a solution to the LCP, $q_{x_0} + My$ gives the stacked paths of the bounded variables in a solution to Problem 5. If, further, either condition of Proposition 10 is satisfied, then this LCP has a unique solution for all $x_0 \in \tilde{X}$, which gives the unique solution to Problem 5, and, for sufficiently large T^*, the finite LCP $\left(q_{x_0}^{(T^*)}, M^{(T^*)}\right)$ has a unique solution $y^{(T^*)}$ for all $x_0 \in \tilde{X}$, where $q_{x_0}^{(T^*)} + M^{(T^*)} y^{(T^*)}$ gives the first T^* periods of the stacked paths of the bounded variables in a solution to Problem 5.

This proposition provides some evidence that the LCP will have a unique solution when it is generated from a dynamic programming problem with a unique solution. In online appendix F, we derive similar results for models with more general constraints and objective functions. The proof of this proposition also showed how one can convert KKT conditions into equations of the form handled by our methods.

3. Applications to New Keynesian models

Brendon, Paustian, and Yates (2015) (henceforth: BPY) consider multiple equilibria in a simple New Keynesian (NK) model with an output growth rate term in the Taylor rule. They show that with sufficiently large reaction to the growth rate, there can be multiple equilibria today, even when the policy rule used to form tomorrow's expectations is held fixed. This is equivalent to the existence of multiple equilibria even when $T = 1$. In the first subsection here, we give an alternative analytic proof of this using our results, and discuss the generalisation to higher T. We go on to consider variants of the BPY model with persistence in shadow nominal interest rates, or price targeting, and show that price targeting produces determinacy.

However, we do not want to give the impression that multiplicity and non-existence are only caused by the central bank responding to the growth rate, or that they are only a problem in carefully constructed theoretical examples. In subsection 3.4, we show that a standard NK model with positive steady-state inflation and a ZLB possesses multiple equilibria in some states, and no solutions in others, even with an entirely standard Taylor rule. We also show that here too price level targeting is sufficient to restore determinacy. Finally, in the last sub-section we show that these conclusions also carry through to the posterior-modes of the Smets and Wouters (2003; 2007) models.

3.1. The simple Brendon, Paustian, and Yates (2015) (BPY) model

The equations of the simple Brendon, Paustian, and Yates (2015) model are as follows:

$$x_{i,t} = \max\{0, 1 - \beta + \alpha_{\Delta y}(x_{y,t} - x_{y,t-1}) + \alpha_\pi x_{\pi,t}\},$$

$$x_{y,t} = \mathbb{E}_t x_{y,t+1} - \frac{1}{\sigma}(x_{i,t} + \beta - 1 - \mathbb{E}_t x_{\pi,t+1}), \qquad x_{\pi,t} = \beta \mathbb{E}_t x_{\pi,t+1} + \gamma x_{y,t},$$

where $x_{i,t}$ is the nominal interest rate, $x_{y,t}$ is the deviation of output from steady-state, $x_{\pi,t}$ is the deviation of inflation from steady-state, and $\beta \in (0,1)$, $\gamma, \sigma, \alpha_{\Delta y} \in (0, \infty)$, $\alpha_\pi \in (1, \infty)$ are parameters. In online appendix G, we prove the following:

Proposition 13 The BPY model is in the form of Problem 2, and satisfies Assumptions 1 and 2. With $T = 1$, $M < 0$ ($M = 0$) if and only if $\alpha_{\Delta y} > \sigma\alpha_\pi$ ($\alpha_{\Delta y} = \sigma\alpha_\pi$).

For a 1×1 matrix, checking the conditions from section 2.3 is trivial. In particular, we have that if $\alpha_{\Delta y} < \sigma\alpha_\pi$, M is a general positive definite, strictly semi-monotone, strictly co-positive, sufficient, P, S matrix; if $\alpha_{\Delta y} \le \sigma\alpha_\pi$, M is a general positive semi-definite, semi-monotone, co-positive, sufficient, P_0, S_0 matrix. Hence, when $T = 1$, if $\alpha_{\Delta y} < \sigma\alpha_\pi$, the model has a unique solution for all q; if $\alpha_{\Delta y} \le \sigma\alpha_\pi$, the model has a unique solution whenever $q > 0$, and at least one solution when $q = 0$. When $\alpha_{\Delta y} > \sigma\alpha_\pi$, M is negative, and so for any positive q, there exists $y > 0$ such that $q + My = 0$, so the model has multiple solutions. I.e. there are solutions that jump to the bound, even when the nominal interest rate would always be positive were there no bound at all.

We illustrate this by adding a shock to the Euler equation, and showing impulse responses for alternative solutions. In particular, we replace the Euler equation with:

$$x_{y,t} = \mathbb{E}_t x_{y,t+1} - \frac{1}{\sigma}(x_{i,t} + \beta - 1 - \mathbb{E}_t x_{\pi,t+1} - (0.01)\varepsilon_t),$$

and take the parameterisation $\sigma = 1$, $\beta = 0.99$, $\gamma = \frac{(1-0.85)(1-\beta(0.85))}{0.85}(2 + \sigma)$, $\rho = 0.5$, following BPY, and we additionally set $\alpha_\pi = 1.5$ and $\alpha_{\Delta y} = 1.6$, to ensure we are in the region with multiple solutions. In Figure 2, we show two alternative solutions to the impulse response to a magnitude 1 shock to ε_t. The solid line in the left plot gives the solution which minimises $\|y\|_\infty$. This solution never hits the bound, and is moderately expansionary. The solid line in the right plot gives the solution which minimises $\|q + My\|_\infty$. (The dotted line in the right plot repeats the left plot, for comparison.) This solution stays at the bound for two periods, and is strongly contractionary, with a magnitude around 100 times larger than the other solution.

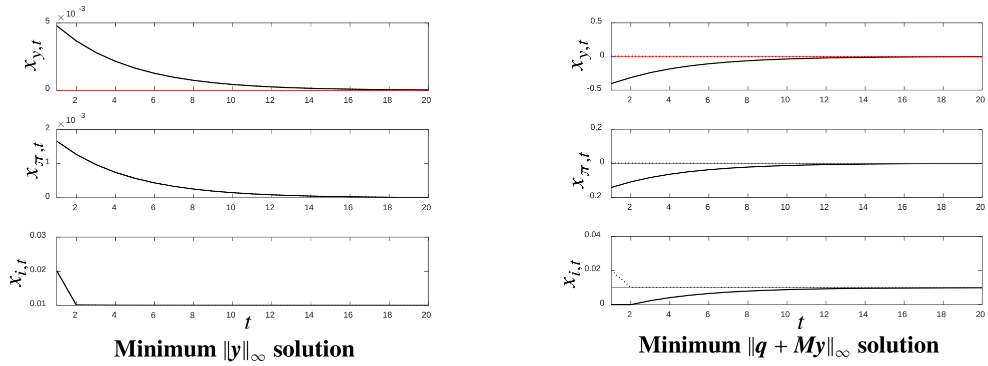


Figure 2: Alternative solutions following a magnitude 1 impulse to ε_t

When $T > 1$, the previous results imply that if $\alpha_{\Delta y} > \sigma\alpha_\pi$, then M is neither P_0, general positive semi-definite, semi-monotone, co-positive, nor sufficient, since the top-left 1×1 principal sub-matrix of M is the same as when $T = 1$. Thus, if anything, when $T > 1$, the parameter region in which there are multiple solutions (when away from the bound or at it) is larger. However, numerical experiments suggest that this parameter region in fact remains the same as T increases, which is unsurprising given the weak persistence of this model. Thus, if we want more interesting results with higher T, we need to consider a model with a stronger persistence mechanism.

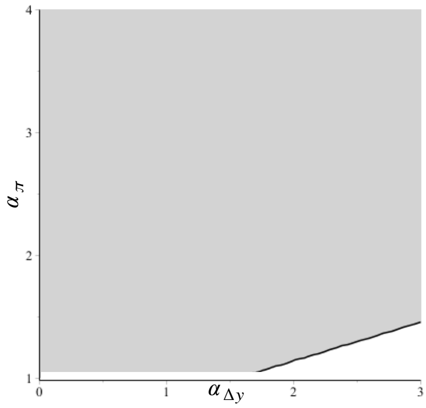
3.2. The BPY model with shadow interest rate persistence

We introduce persistence in the shadow interest rate by replacing the previous Taylor rule with $x_{i,t} = \max\{0, x_{d,t}\}$, where $x_{d,t}$, the shadow nominal interest rate is given by:

$$x_{d,t} = (1 - \rho)\left(1 - \beta + \alpha_{\Delta y}\left(x_{y,t} - x_{y,t-1}\right) + \alpha_\pi x_{\pi,t}\right) + \rho x_{d,t-1}.$$

It is easy to verify that this may be put in the form of Problem 2, and that with $\beta \in (0,1)$, $\gamma, \sigma, \alpha_{\Delta y} \in (0, \infty)$, $\alpha_\pi \in (1, \infty)$, $\rho \in (-1,1)$, Assumption 2 is satisfied. For our numerical exercise, we again set $\sigma = 1$, $\beta = 0.99$, $\gamma = \frac{(1-0.85)(1-\beta(0.85))}{0.85}(2 + \sigma)$, $\rho = 0.5$, following BPY.

In Figure 3, we plot the regions in $(\alpha_{\Delta y}, \alpha_\pi)$ space in which M is a P-matrix (P_0-matrix) when $T = 2$ or $T = 4$. For this model, these correspond to the regions in which M is strictly semi-monotone (semi-monotone). As may be seen, in the smaller T case, the P-matrix region is much larger. This relationship appears to continue to hold for both larger and smaller T, with the equivalent $T = 1$ plot being almost entirely shaded, and the large T plot apparently tending to the equivalent plot from the model without monetary policy persistence. Intuitively, the persistence in the shadow nominal interest rate dampens the immediate response of nominal interest rates to inflation and output growth, making it harder to induce a zero lower bound episode over short-horizons.



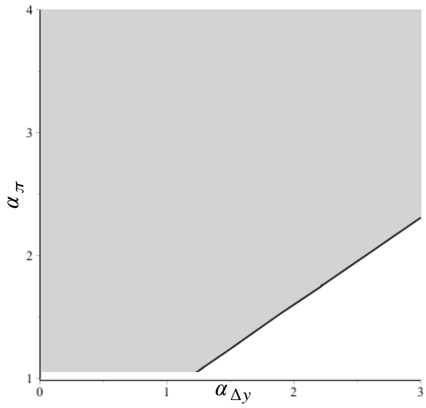


Figure 3: Regions in which M is a P-matrix (shaded grey) or a P_0-matrix (shaded grey, plus the black line), when $T = 2$ (left) or $T = 4$ (right).

Further evidence that the long-horizon behaviour is the same as in the model without persistence is provided by the fact that with $\alpha_\pi = 1.5$ and $\alpha_{\Delta y} = 1.05$, [14] then M is a P-matrix, and from Proposition 4 we have that $\varsigma > 6.131 \times 10^{-8}$, so M is an S-matrix for all sufficiently large T. Furthermore, with $\alpha_\pi = 1.5$ and $\alpha_{\Delta y} = 1.51$, then with $T = 200$, M is not an S-matrix, [15] and from Proposition 4, $\varsigma \le 0 + \text{numerical error}$, providing strong numerical evidence that for all sufficiently large T, the LCP (q, M) is not feasible for some q, and hence that the model does not always possess a solution.

3.3. The BPY model with price targeting

An alternative way to introduce persistence to the shadow interest rate is to set:

$$\begin{aligned} x_{d,t} &= (1-\rho)\left(1 - \beta + \frac{\alpha_{\Delta y}}{1-\rho}\left(x_{y,t} - x_{y,t-1}\right) + \frac{\alpha_\pi}{1-\rho} x_{\pi,t}\right) + \rho x_{d,t-1} \\ &= (1-\rho)(1-\beta) + \left(\alpha_{\Delta y}\left(x_{y,t} - x_{y,t-1}\right) + \alpha_\pi x_{\pi,t}\right) + \rho x_{d,t-1}, \end{aligned}$$

which is as before apart from a missing $(1-\rho)$ multiplying the second bracketed term. In the limit as $\rho \to 1$, this tends to:

$$x_{d,t} = 1 - \beta + \alpha_{\Delta y} x_{y,t} + \alpha_\pi x_{p,t}$$

where $x_{p,t}$ is the price level, so $x_{\pi,t} = x_{p,t} - x_{p,t-1}$. This is a level targeting rule, with nominal GDP targeting as a special case with $\alpha_{\Delta y} = \alpha_\pi$. Note that the omission of the $(1-\rho)$ coefficient on $\alpha_{\Delta y}$ and α_π is akin to having a "true" response to output growth of $\frac{\alpha_{\Delta y}}{1-\rho}$ and a "true" response to inflation of $\frac{\alpha_\pi}{1-\rho}$, so in the limit as $\rho \to 1$, we effectively have an infinitely strong response to these quantities. It turns out that this is sufficient to produce determinacy for all $\alpha_{\Delta y}, \alpha_\pi \in (0, \infty)$.

[14] Results for larger $\alpha_{\Delta y}$ were impossible due to numerical errors.

[15] This was verified a second way by checking that $-M'$ was an S_0-matrix, as discussed in footnote 9.

In particular, given the model:

$$x_{i,t} = \max\{0, 1-\beta + \alpha_{\Delta y} x_{y,t} + \alpha_\pi x_{p,t}\},$$
$$x_{y,t} = \mathbb{E}_t x_{y,t+1} - \frac{1}{\sigma}\left(x_{i,t} + \beta - 1 - \mathbb{E}_t x_{p,t+1} + x_{p,t}\right),$$
$$x_{p,t} - x_{p,t-1} = \beta \mathbb{E}_t x_{p,t+1} - \beta x_{p,t} + \gamma x_{y,t},$$

we prove in online appendix H that the following proposition holds:

Proposition 14 The BPY model with price targeting is in the form of Problem 2, and satisfies Assumptions 1 and 2. With $T = 1$, $M > 0$ for all $\alpha_\pi \in (0,\infty)$, $\alpha_{\Delta y} \in [0,\infty)$.

Furthermore, with $\sigma = 1$, $\beta = 0.99$, $\gamma = \frac{(1-0.85)(1-\beta(0.85))}{0.85}(2+\sigma)$, as before, and $\alpha_{\Delta y} = 1$, $\alpha_\pi = 1$, if we check our lower bound on ς with $T = 20$, we find that $\varsigma > 0.042$. Hence, this model is always feasible for any sufficiently large T. Given that $d_0 > 0$ for this model, and that for $T = 20$, M is a P-matrix, this is strongly suggestive of the existence of a unique solution for any q and for arbitrarily large T.

3.4. The linearized Fernández-Villaverde et al. (2012) model

The discussion of BPY might lead one to believe that multiplicity and non-existence is solely a consequence of overly aggressive monetary responses to output growth, and overly weak monetary responses to inflation. However, it turns out that in basic New Keynesian models with positive inflation in steady-state, and hence price dispersion, even without any monetary response to output growth, and even with extremely aggressive monetary responses to inflation, there are still multiple equilibria in some states of the world, and no solutions in others. Price level targeting is again sufficient to fix these problems though.

We show these results in the Fernández-Villaverde et al. (2012) model, which is a basic non-linear New Keynesian model without capital or price indexation of non-resetting firms, but featuring (non-valued) government spending and steady-state inflation (and hence price-dispersion). We refer the reader to the original paper for the model's equations. After substitutions, the model has four non-linear equations which are functions of gross inflation, labour supply, price dispersion and an auxiliary variable introduced from the firms' price-setting first order condition. Of these variables, only price dispersion enters with a lag. We linearize[16] the model around its steady-state, and then reintroduce the "max" operator which linearization removed from the Taylor rule. All parameters are set to the values given in Fernández-Villaverde et al. (2012). There

[16] Prior to linearization, we first transform the model's variables so that the transformed variables may take values on the entire real line. I.e. we work with the logarithms of labour supply, price dispersion and the auxiliary variable. For inflation, we note that inflation is always less than $\theta^{\frac{1}{1-\varepsilon}}$ (in the notation of Fernández-Villaverde et al. (2012)). Thus we work with a logit transformation of inflation over $\theta^{\frac{1}{1-\varepsilon}}$. This is generally more accurate than working with the logarithm of inflation.

is no term featuring output growth in the Taylor rule, so any multiplicity or non-existence in this model cannot be a consequence of the mechanism highlighted by BPY.

For this model, numerical calculations reveal that with $T \le 14$, M is a P-matrix. However, with $T = 15$, M is neither a P nor an S matrix, and thus there are certainly some states of the world in which the model has multiple solutions, and others in which it has no solution at all.[17] This also implies that M is not a P-matrix for all larger T. Furthermore, with $T = 1000$, our upper bound on ς from Proposition 4 implies that $\varsigma \le 0$ + numerical error, providing evidence that M is not an S-matrix for large T either.[18]

However, if we replace inflation in the monetary rule with the price level relative to its linear trend, which evolves according to:

$$x_{p,t} = x_{p,t-1} + x_{\pi,t} - x_{\pi}, \tag{16}$$

then with $T = 200$, we have that M is an S-matrix, and the lower bound from Proposition 4 implies that $\varsigma > 0.003$, and hence that for all sufficiently large T, M is an S-matrix, so there is always a feasible solution.

3.5. The Smets and Wouters (2003) and Smets and Wouters (2007) models

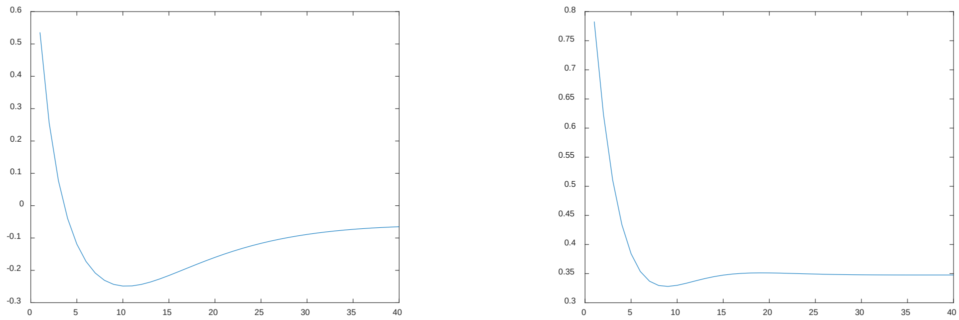
Smets and Wouters (2003) and Smets and Wouters (2007) are the canonical medium-scale linear DSGE models, featuring assorted shocks, habits, price and wage indexation, capital (with adjustment costs), (costly) variable utilisation and quite general monetary policy reaction functions. The former model is estimated on Euro area data, while the latter is estimated on US data. The latter model also contains trend growth (permitting its estimation on non-detrended data), and a slightly more general aggregator across industries. However, overall, they are quite similar models, and any differences in their behaviour chiefly stems from differences in the estimated parameters. Since both models are incredibly well known in the literature, we omit their equations here, referring the reader to the original papers for further details.

To assess the likelihood of multiple equilibria at or away from the zero lower bound, we augment each model with a ZLB on nominal interest rates, and evaluate the properties of each model's M matrix with large T, at the estimated posterior-modes from the original papers. Note that we do not introduce an auxiliary for shadow nominal interest rates, so the monetary rules take the form of $x_{r,t} = \max\{0, (1-\rho_r)(\cdots) + \rho_r x_{r,t-1} + \cdots\}$, in both cases.

[17] That New Keynesian models might have no solution at all in some states of the world has also been discussed by Basu and Bundick (2015), though their mechanism only applies in the stochastic model.

[18] Since these results depend on the presence of the endogenous state, price dispersion, they are not directly related to the results of Davig and Leeper (2007). Further differences include the endogeneity of ZLB episodes here, and the fact that we are not making any restrictions on the solution space, which they do, as observed by Farmer, Waggoner, and Zha (2010).

As shown in Lemma 4, if the diagonal of the M matrix ever goes negative, then the M matrix cannot be general positive semi-definite, semi-monotone, sufficient, P_0 or copositive, and hence the model will sometimes have multiple solutions even when away from the zero lower bound (i.e. for some strictly positive q). In Figure 4, we plot the diagonal of the M matrix for each model in turn,[19] i.e. the impact on nominal interest rates in period t of news in period 1 that a positive, magnitude one shock will hit nominal interest rates in period t. Immediately, we see that while in the US model, these impacts remain positive at all horizons, in the Euro area model, these impacts turn negative after just a few periods, and remain so at least up to period 40. Therefore, in the ZLB augmented Smets and Wouters (2003) model, there is not always a unique equilibrium. Furthermore, there are sequences of predicted future shocks (with positive density) for which the model without the ZLB would always feature positive interest rates, but for which the model with the ZLB could hit zero.



The Smets and Wouters (2003) model The Smets and Wouters (2007) model

Figure 4: The diagonals of the *M* matrices for the Smets and Wouters (2003; 2007) models

It remains for us to assess whether M is a $P_{(0)}$-matrix or (strictly) semi-monotone for the Smets and Wouters (2007) model. Numerical calculations reveal that for $T < 9$, M is a P-matrix, and hence is strictly semi-monotone. However, with $T \geq 9$, M contains a 6×6 principal sub-matrix (with indices 1,2,4,6,7,9) with negative determinant, which is neither an S nor an S_0-matrix. Thus, for $T \geq 9$, M is not a $P_{(0)}$-matrix or (strictly) semi-monotone, and hence this model also has multiple equilibria, even when away from the bound. Given that the US has been at the ZLB for over eight years, that T ought to be greater than eight quarters seems uncontroversial. Hence, in both the Euro area and the US, we ought to take seriously the possibility that the existence of the ZLB produces non-uniqueness. Furthermore, it turns out that for neither model is M an S-matrix even with $T = 1000$, and thus for both models there are some $q \in \mathbb{R}^{1000}$ for

[19] The MOD files for the Smets and Wouters (2003) model were derived from the Macro Model Database (Wieland et al. 2012). The MOD files for the model were derived from files provided by Johannes Pfeifer here: http://goo.gl/CP53x5

which no solution exists. This is strongly suggestive of non-existence for some q even for arbitrarily large T. While placing a larger coefficient on inflation in the Taylor rule can make the Euro area picture more like the US one, with a strictly positive diagonal to the M matrix, even with incredibly large coefficients, M remained a non-P-matrix.

Alternatively, suppose we replace the monetary rule in both models by:

$$x_{r,t} = \max\{0, (1-\rho_r)(x_{y,t} + x_{p,t}) + \rho_r x_{r,t-1}\}$$

where ρ_r is as in the respective original model, where the price level $x_{p,t}$ again evolves according to equation (16), and where $x_{y,t}$ is output relative to its linear trend. Then, for both models, for all T tested, M was a P-matrix, and for the Euro area model we have that $\varsigma > 3 \times 10^{-7}$ and for the US model we have that $\varsigma > 0.002$. As one would expect, this result is also robust to departures from equal, unit, coefficients. Thus, price level targeting again appears to be sufficient for determinacy in the presence of the ZLB.

4. Conclusion

This paper provides the first theoretical results on existence and uniqueness for otherwise linear models with occasionally binding constraints. As such, it may be thought of as doing for models with occasionally binding constraints what Blanchard and Kahn (1980) did for linear models.

We provided necessary and sufficient conditions for the existence of a unique equilibrium, as well as such conditions for uniqueness when away from the bound. In our application to New Keynesian models, we showed that these conditions were violated in entirely standard models, rather than being an artefact of strange policy rules as one might have inferred from the results of Brendon, Paustian, and Yates (2015). In the presence of multiplicity, there is the potential for additional endogenous volatility from sunspots, so the welfare benefits of avoiding multiplicity may be substantial. Additionally, as we saw in Figure 2, the additional equilibria may feature huge drops in output, giving further welfare reasons for their avoidance. The possibility of self-fulfilling jumps and returns from the ZLB also gives an alternative rationale for the neo-Fisherian view that argues that raising interest rates may raise inflation at the ZLB.[20]

Luckily, our results suggest that a determinate equilibrium may be produced in standard New Keynesian models if the central bank switches to targeting the price level, rather than the inflation rate. This provides an additional argument for price level targeting in the presence of a zero lower bound to those made by Basu and Bundick (2015) and Coibion, Gorodnichenko, and Wieland (2012). Indeed, it is possible that Coibion, Gorodnichenko, and Wieland's results on the welfare benefits of price level

[20] Theoretical and empirical evidence for this view is presented in Cochrane (2015).

targeting were actually driven by having inadvertently selected one of the worse equilibria under inflation targeting, since they use a solution algorithm for the otherwise linear model which gives no guarantees on the returned equilibrium.

In addition, we provided conditions for existence of any solution that converges to the "good" steady-state, and showed that under inflation targeting, standard New Keynesian models again failed to satisfy these conditions in some states of the world. Whereas the literature started by Benhabib, Schmitt-Grohé, and Uribe (2001a; 2001b) showed that the existence of a "bad" steady-state may imply additional volatility if agents long-run beliefs are not pinned down by the inflation target, here we showed that in some states of the world, under inflation targeting there is no way for the economy to converge to the "good" steady-state. This in turn implies that agents cannot place prior certainty on converging to the "good" steady-state, thus rationalising the beliefs required to get the kind of global multiplicity at the zero lower bound that these and other authors have focussed on. Once again though, we showed that price level targeting is sufficient to restore existence and determinacy.

5. References

Adjemian, Stéphane, and Michel Juillard. 2013. 'Stochastic Extended Path Approach'.

Aliprantis, C.D., and K. Border. 2013. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Studies in Economic Theory. Springer Berlin Heidelberg.

Anderson, Gary. 2015. 'A Solution Strategy for Occasionally Binding Constraints'. Unpublished.

Aruoba, S. Borağan, Pablo Cuba-Borda, and Frank Schorfheide. 2014. *Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries*. Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Basu, Susanto, and Brent Bundick. 2015. 'Endogenous Volatility at the Zero Lower Bound: Implications for Stabilization Policy'. *National Bureau of Economic Research Working Paper Series* No. 21838.

Beaudry, Paul, and Franck Portier. 2006. 'Stock Prices, News, and Economic Fluctuations'. *American Economic Review* 96 (4): 1293–1307.

Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe. 2001a. 'Monetary Policy and Multiple Equilibria'. *American Economic Review* 91 (1). American Economic Review: 167–186.

———. 2001b. 'The Perils of Taylor Rules'. *Journal of Economic Theory* 96 (1-2). Journal of Economic Theory: 40–69.

Benigno, Pierpaolo, and Michael Woodford. 2012. 'Linear-Quadratic Approximation of Optimal Policy Problems'. *Journal of Economic Theory* 147 (1). Journal of Economic Theory: 1–42.

Blanchard, Olivier Jean, and Charles M. Kahn. 1980. 'The Solution of Linear Difference Models under Rational Expectations'. *Econometrica* 48 (5): 1305–1311.

Bodenstein, Martin, Luca Guerrieri, and Christopher J. Gust. 2013. 'Oil Shocks and the Zero Bound on Nominal Interest Rates'. *Journal of International Money and Finance* 32 (February): 941–967.

Brendon, Charles, Matthias Paustian, and Tony Yates. 2015. 'Self-Fulfilling Recessions at the Zero Lower Bound'. *Unpublished*.

Cochrane, John H. 2015. 'Do Higher Interest Rates Raise or Lower Inflation?' Unpublished.

Coibion, Olivier, Yuriy Gorodnichenko, and Johannes Wieland. 2012. 'The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?' *The Review of Economic Studies* 79 (4) (October 1): 1371–1406.

Cottle, Richard W. 2009. 'Linear Complementarity Problem'. In *Encyclopedia of Optimization*, edited by Christodoulos A. Floudas and Panos M. Pardalos, 1873–1878. Springer US.

Cottle, Richard W., Jong-Shi Pang, and Richard E. Stone. 2009a. '3. Existence and Multiplicity'. In *The Linear Complementarity Problem*, 137–223. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.

———. 2009b. '2. Background'. In *The Linear Complementarity Problem*, 43–136. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.

Davig, Troy, and Eric M. Leeper. 2007. 'Generalizing the Taylor Principle'. *American Economic Review* 97 (3): 607–635.

Fang, Li. 1989. 'On the Spectra of P- and P0-Matrices'. *Linear Algebra and Its Applications* 119 (July): 1–25.

Farmer, Roger E. A., Daniel F. Waggoner, and Tao Zha. 2010. 'Generalizing the Taylor Principle: Comment'. *American Economic Review* 100 (1): 608–17.

Fernández-Villaverde, Jesús, Grey Gordon, Pablo Guerrón-Quintana, and Juan F. Rubio-Ramírez. 2012. 'Nonlinear Adventures at the Zero Lower Bound'. NBER Working Papers.

Gavin, William T., Benjamin D. Keen, Alexander W. Richter, and Nathaniel A. Throckmorton. 2015. 'The Zero Lower Bound, the Dual Mandate, and Unconventional Dynamics'. *Journal of Economic Dynamics and Control* 55 (June): 14–38.

Hebden, James, Jesper Lindé, and Lars Svensson. 2011. 'Optimal Monetary Policy in the Hybrid New Keynesian Model under the Zero Lower Bound'. Unpublished.

Holden, Tom. 2010. *Products, Patents and Productivity Persistence: A DSGE Model of Endogenous Growth*. Economics Series Working Paper. University of Oxford, Department of Economics.

———. 2016. 'Existence and Uniqueness of Solutions to Dynamic Models with Occasionally Binding Constraints.' Unpublished.

Holden, Tom, and Michael Paetz. 2012. 'Efficient Simulation of DSGE Models with Inequality Constraints'. School of Economics Discussion Papers.

Kamihigashi, Takashi. 2014. 'Elementary Results on Solutions to the Bellman Equation of Dynamic Programming: Existence, Uniqueness, and Convergence'. *Economic Theory* 56 (2) (June 1): 251–273.

Kamihigashi, Takashi, and Santanu Roy. 2003. *A Nonsmooth, Nonconvex Model of Optimal Growth*. Discussion Paper Series. Research Institute for Economics & Business Administration, Kobe University.

Levine, Paul, Joseph Pearlman, and Richard Pierse. 2008. 'Linear-Quadratic Approximation, External Habit and Targeting Rules'. *Journal of Economic Dynamics and Control* 32 (10). Journal of Economic Dynamics and Control: 3315–3349.

Mertens, Karel RSM, and Morten O. Ravn. 2014. 'Fiscal Policy in an Expectations-Driven Liquidity Trap'. *The Review of Economic Studies*: rdu016.

Richter, Alexander W., and Nathaniel A. Throckmorton. 2014. 'The Zero Lower Bound: Frequency, Duration, and Numerical Convergence'. *The B.E. Journal of Macroeconomics* 15 (1): 157.

Roos, C., T. Terlaky, and J. P. Vial. 2006. *Interior Point Methods for Linear Optimization*. Springer.

Rotemberg, J. 1982. 'Monopolistic Price Adjustment and Aggregate Output'. *Review of Economic Studies* 49: 517–531.

Samelson, Hans, Robert M. Thrall, and Oscar Wesler. 1958. 'A Partition Theorem for Euclidean N-Space'. *Proceedings of the American Mathematical Society* 9 (5): 805–807.

Schmitt-Grohé, Stephanie, and Martín Uribe. 2012. *The Making Of A Great Contraction With A Liquidity Trap and A Jobless Recovery*. National Bureau of Economic Research, Inc.

Sims, ChristopherA. 2002. 'Solving Linear Rational Expectations Models'. *Computational Economics* 20 (1-2) (October 1): 1–20.

Smets, Frank, and Rafael Wouters. 2003. 'An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area'. *Journal of the European Economic Association* 1 (5): 1123–1175.

———. 2007. 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach'. *American Economic Review* 97 (3) (June): 586–606.

Stokey, N., R. J. Lucas, and E. Prescott. 1989. *Recursive Methods in Economic Dynamics*. Harvard University Press.

Trefethen, L.N., and M. Embree. 2005. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press.

Väliaho, Hannu. 1986. 'Criteria for Copositive Matrices'. *Linear Algebra and Its Applications* 81 (September 1): 19–34.

Wieland, Volker, Tobias Cwik, Gernot J. Müller, Sebastian Schmidt, and Maik Wolters. 2012. 'A New Comparative Approach to Macroeconomic Modeling and Policy Analysis'. *The Great Recession: Motivation for Re-Thinking Paradigms in Macroeconomic Modeling* 83 (3) (August): 523–541.

Wright, Thomas G, and Lloyd N Trefethen. 2001. 'Large-Scale Computation of Pseudospectra Using ARPACK and Eigs'. *SIAM Journal on Scientific Computing* 23 (2): 591–605.

Xu, Song. 1993. 'Notes on Sufficient Matrices'. *Linear Algebra and Its Applications* 191 (September 15): 1–13.

Online Appendices to: “Existence and uniqueness of solutions to dynamic models with occasionally binding constraints.”

Tom D. Holden, School of Economics, University of Surrey

A. Construction of a static model with no dynamic solution in some states

Consider the model:

$$a_t = \max\{0, b_t\}, \qquad a_t = 1 - c_t, \qquad c_t = a_t - b_t.$$

The model has steady-state $a = b = 1$, $c = 0$. Furthermore, in the model’s Problem 3 type equivalent, in which for $t \in \mathbb{N}^+$:

$$a_t = \begin{cases} b_t + y_{t,0} & \text{if } t \le T \\ b_t & \text{if } t > T \end{cases},$$

where $y_{\cdot,\cdot}$ is defined as in Problem 3, we have that:

$$c_t = \begin{cases} y_{t,0} & \text{if } t \le T \\ 0 & \text{if } t > T \end{cases},$$

so:

$$b_t = \begin{cases} 1 - 2y_{t,0} & \text{if } t \le T \\ 1 & \text{if } t > T \end{cases},$$

implying:

$$a_t = \begin{cases} 1 - y_{t,0} & \text{if } t \le T \\ 1 & \text{if } t > T \end{cases}.$$

thus, $M = -I$ for this model.

B. Proof of sufficient conditions for feasibility with $T = \infty$

First, define $G \coloneqq -C(B + CF)^{-1}$, and note that if L is the lag (right-shift) operator, the model from Problem 1 can be written as:

$$L^{-1}(ALL + BL + C)(x - \mu) = 0.$$

Furthermore, by the definitions of F and G:

$$(L - G)(B + CF)(I - FL) = ALL + BL + C,$$

so the stability of the model from Problem 1 is determined by the solutions for $z \in \mathbb{C}$ of the polynomial:

$$0 = \det(Az^2 + Bz + C) = \det(Iz - G)\det(B + CF)\det(I - Fz).$$

Now by Assumption 1, all of the roots of $\det(I - Fz)$ are strictly outside of the unit circle, and all of the other roots of $\det(Az^2 + Bz + C)$ are weakly inside the unit circle (else there would be indeterminacy), thus, all of the roots of $\det(Iz - G)$ are weakly inside the unit circle. Therefore, if we write $\rho_{\mathscr{M}}$ for the spectral radius of some matrix $\mathscr{M}$, then, by this discussion and Assumption 2, $\rho_G < 1$.

Next, let $s_t^*, x_t^* \in \mathbb{R}^{n\times\mathbb{N}^+}$ be such that for any $y \in \mathbb{R}^{\mathbb{N}^+}$, the k^{th} columns of $s_t^* y$ and $x_t^* y$ give the value of s_t and x_t following a magnitude 1 news shock at horizon k, i.e. when $x_0 = \mu$ and y_0 is the k^{th} row of $I_{\mathbb{N}^+\times\mathbb{N}^+}$. Then:

$$\begin{aligned}s_t^* &= -(B+CF)^{-1}[I_{\cdot,1}I_{t,1:\infty} + GI_{\cdot,1}I_{t+1,1:\infty} + G^2 I_{\cdot,1}I_{t+2,1:\infty} + \cdots] \\ &= -(B+CF)^{-1}\sum_{k=0}^{\infty}(GL)^k I_{\cdot,1}I_{t,1:\infty} \\ &= -(B+CF)^{-1}(I-GL)^{-1}I_{\cdot,1}I_{t,1:\infty},\end{aligned}$$

where the infinite sums are well defined as $\rho_G < 1$, and where $I_{t,1:\infty} \in \mathbb{R}^{1\times\mathbb{N}^+}$ is a row vector with zeros everywhere except position t where there is a 1. Thus:

$$s_t^* = \begin{bmatrix} 0_{n\times(t-1)} & s_1^* \end{bmatrix} = L^{t-1}s_1^*.$$

Furthermore,

$$(x_t^* - \mu^*) = \sum_{j=1}^{t} F^{t-j}s_k^* = \sum_{j=1}^{t} F^{t-j}L^{j-1}s_1^*,$$

i.e.:

$$(x_t^* - \mu^*)_{\cdot,k} = \sum_{j=1}^{t} F^{t-j}s_{1,\cdot,k+1-j}^* = -\sum_{j=1}^{\min\{t,k\}} F^{t-j}(B+CF)^{-1}G^{k-j}I_{\cdot,1},$$

and so the k^{th} offset diagonal of M ($k \in \mathbb{Z}$) is given by the first row of the k^{th} column of:

$$L^{-t}(x_t^* - \mu^*) = L^{-1}\sum_{j=1}^{t}(FL^{-1})^{t-j}s_1^* = L^{-1}\sum_{j=0}^{t-1}(FL^{-1})^j s_1^*,$$

where we abuse notation slightly by allowing L^{-1} to give a result with indices in $\mathbb{Z}$ rather than $\mathbb{N}^+$, with padding by zeros. Consequently, for all $k \in \mathbb{N}^+$, $M_{t,k} = \mathrm{O}(t^n\rho_F^t)$, as $t \to \infty$, for all $t \in \mathbb{N}^+$, $M_{t,k} = \mathrm{O}\big(t^n\rho_G^k\big)$, as $k \to \infty$, and for all $k \in \mathbb{Z}$, $M_{t,t+k} - \lim_{\tau\to\infty} M_{\tau,\tau+k} = \mathrm{O}(t^{n-1}(\rho_F\rho_G)^t)$, as $t \to \infty$. Hence,

$$\sup_{y\in[0,1]^{\mathbb{N}^+}}\inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y$$

exists and is well defined, and so:

$$\varsigma = \sup_{\substack{y\in[0,1]^{\mathbb{N}^+} \\ \exists T\in\mathbb{N} \text{ s.t. } \forall t>T, y_t=0}}\inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y = \sup_{y\in[0,1]^{\mathbb{N}^+}}\inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y,$$

since every point in $[0,1]^{\mathbb{N}^+}$ is a limit (under the supremum norm) of a sequence of points in the set:

$$\big\{y \in [0,1]^{\mathbb{N}^+} \big| \exists T \in \mathbb{N} \text{ s.t. } \forall t > T, y_t = 0\big\}.$$

Thus, we just need to provide conditions under which $\sup_{y\in[0,1]^{\mathbb{N}^+}}\inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y > 0$.

To produce such conditions, we need constructive bounds on M, even if they have slightly worse convergence rates. For any matrix, $\mathscr{M} \in \mathbb{R}^{n \times n}$ with $\rho_{\mathscr{M}} < 1$, and any $\phi \in (\rho_{\mathscr{M}}, 1)$, let:

$$\mathscr{C}_{\mathscr{M},\phi} \coloneqq \sup_{k \in \mathbb{N}} \|(\mathscr{M}\phi^{-1})^k\|_2.$$

Furthermore, for any matrix, $\mathscr{M} \in \mathbb{R}^{n \times n}$ with $\rho_{\mathscr{M}} < 1$, and any $\epsilon > 0$, let:

$$\rho_{\mathscr{M},\epsilon} \coloneqq \max\{|z| | z \in \mathbb{C}, \sigma_{\min}(\mathscr{M} - zI) = \epsilon\},$$

where $\sigma_{\min}(\mathscr{M} - zI)$ is the minimum singular value of $\mathscr{M} - zI$, and let $\epsilon^*(\mathscr{M}) \in (0, \infty]$ solve:

$$\rho_{\mathscr{M},\epsilon} = 1.$$

(This has a solution in $(0, \infty]$ by continuity as $\rho_{\mathscr{M}} < 1$.) Then, by Theorem 16.2 of Trefethen and Embree (2005), for any $K \in \mathbb{N}$ and $k > K$:

$$\|(\mathscr{M}\phi^{-1})^k\|_2 \le \|(\mathscr{M}\phi^{-1})^K\|_2 \|(\mathscr{M}\phi^{-1})^{k-K}\|_2 \le \frac{\|(\mathscr{M}\phi^{-1})^K\|_2}{\epsilon^*(\mathscr{M}\phi^{-1})}.$$

Now, $\|(\mathscr{M}\phi^{-1})^K\|_2 \to 0$ as $K \to \infty$, hence, there exists some $K \in \mathbb{N}$ such that:

$$\sup_{k=0,\dots,K} \|(\mathscr{M}\phi^{-1})^k\|_2 \ge \frac{\|(\mathscr{M}\phi^{-1})^K\|_2}{\epsilon^*(\mathscr{M}\phi^{-1})} \ge \sup_{k>K} \|(\mathscr{M}\phi^{-1})^k\|_2,$$

meaning $\mathscr{C}_{\mathscr{M},\phi} = \sup_{k=0,\dots,K} \|(\mathscr{M}\phi^{-1})^k\|_2$. The quantity $\rho_{\mathscr{M},\epsilon}$ (and hence $\epsilon^*(\mathscr{M})$) may be efficiently computed using the methods described by Wright and Trefethen (2001), and implemented in their EigTool toolkit[21]. Thus, $\mathscr{C}_{\mathscr{M},\phi}$ may be calculated in finitely many operations by iterating over $K \in \mathbb{N}$ until a K is found which satisfies:

$$\sup_{k=0,\dots,K} \|(\mathscr{M}\phi^{-1})^k\|_2 \ge \frac{\|(\mathscr{M}\phi^{-1})^K\|_2}{\epsilon^*(\mathscr{M}\phi^{-1})}.$$

From the definition of $\mathscr{C}_{\mathscr{M},\phi}$, we have that for any $k \in \mathbb{N}$ and any $\phi \in (\rho_{\mathscr{M}}, 1)$:

$$\|\mathscr{M}^k\|_2 \le \mathscr{C}_{\mathscr{M},\phi} \phi^k.$$

Now, fix $\phi_F \in (\rho_F, 1)$ and $\phi_G \in (\rho_G, 1)$,[22] and define:

$$\mathscr{D}_{\phi_F,\phi_G} \coloneqq \mathscr{C}_{F,\phi_F} \mathscr{C}_{G,\phi_F} \|(B + CF)^{-1}\|_2,$$

then, for all $t, k \in \mathbb{N}^+$:

$$\begin{aligned}|M_{t,k}| = |(x_t^* - \mu^*)_{1,k}| \le \|(x_t^* - \mu^*)_{\cdot,k}\|_2 &\le \sum_{j=1}^{\min\{t,k\}} \|F^{t-j}\|_2 \|(B + CF)^{-1}\|_2 \|G^{k-j}\|_2 \\ \le \mathscr{D}_{\phi_F,\phi_G} \sum_{j=1}^{\min\{t,k\}} \phi_F^{t-j} \phi_G^{k-j} &= \mathscr{D}_{\phi_F,\phi_G} \phi_F^t \phi_G^k \frac{(\phi_F \phi_G)^{-\min\{t,k\}} - 1}{1 - \phi_F \phi_G}.\end{aligned}$$

Additionally, for all $t \in \mathbb{N}^+$, $k \in \mathbb{Z}$:

$$\left|M_{t,t+k} - \lim_{\tau \to \infty} M_{\tau,\tau+k}\right| = \left|\left(L^{-t}(x_t^* - \mu^*)\right)_{1,k} - \left(\lim_{\tau \to \infty} L^{-t}(x_t^* - \mu^*)\right)_{1,k}\right|$$

[21] This toolkit is available from https://github.com/eigtool/eigtool, and is included in dynareOBC.

[22] In practice, we try a grid of values, as it is problem dependent whether high ϕ_F and low $\mathcal{K}(\mathcal{M}\phi^{-1})$ is preferable to low ϕ_F and high $\mathcal{K}(\mathcal{M}\phi^{-1})$.

$$
\begin{aligned}
&\leq \left\| \left(L^{-1} \sum_{j=0}^{t-1} (FL^{-1})^j s_1^* - L^{-1} \sum_{j=0}^{\infty} (FL^{-1})^j s_1^* \right)_{\cdot,k} \right\|_2 \\
&= \left\| \left(\sum_{j=\max\{t,-k\}}^{\infty} F^j s_{1,\cdot,j+k+1}^* \right)_{\cdot,0} \right\|_2 \\
&= \left\| \sum_{j=\max\{t,-k\}}^{\infty} F^j (B+CF)^{-1} G^{j+k} I_{\cdot,1} \right\|_2 \\
&\leq \sum_{j=\max\{t,-k\}}^{\infty} \|F^j\|_2 \|(B+CF)^{-1}\|_2 \|G^{j+k}\|_2 \\
&\leq \mathscr{D}_{\phi_F,\phi_G} \sum_{j=\max\{t,-k\}}^{\infty} \phi_F^j \phi_G^{j+k} = \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^{\max\{t,-k\}} \phi_G^{\max\{0,t+k\}}}{1-\phi_F\phi_G},
\end{aligned}
$$

so, for all $t, k \in \mathbb{N}^+$:

$$
\left| M_{t,k} - \lim_{\tau\to\infty} M_{\tau,\tau+k-t} \right| \leq \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G}.
$$

To evaluate $\lim_{\tau\to\infty} M_{\tau,\tau+k-t}$, note that this limit is the top element from the $(k-t)^{\text{th}}$ column of:

$$
\begin{aligned}
d &\coloneqq \lim_{\tau\to\infty} L^{-\tau}(x_\tau^* - \mu^*) = L^{-1}(I - FL^{-1})^{-1} s_1^* \\
&= -(I - FL^{-1})^{-1}(B+CF)^{-1}(I - GL)^{-1} I_{\cdot,1} I_{0,-\infty:\infty},
\end{aligned}
$$

where $I_{0,-\infty:\infty} \in \mathbb{R}^{1\times\mathbb{Z}}$ is zero everywhere apart from index 0 where it equals 1. Hence, by the definitions of F and G:

$$
AL^{-1}d + Bd + CLd = -I_{\cdot,1} I_{0,-\infty:\infty}.
$$

In other words, if we write d_k in place of $d_{\cdot,k}$ for convenience, then, for all $k \in \mathbb{Z}$:

$$
Ad_{k+1} + Bd_k + Cd_{k-1} = -\begin{cases} I_{\cdot,1} & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}
$$

I.e. the homogeneous part of the difference equation for d_{-t} is identical to that of $x_t - \mu$. The time reversal here is intuitive since we are indexing diagonals such that indices increase as we move up and to the right in M, but time is increasing as we move down in M.

It turns out that exploiting the possibility of reversing time is the key to easy evaluating d_k. First, note that for $k < 0$, it must be the case that $d_k = Fd_{k+1}$, since the shock has already "occurred" (remember, that we are going forwards in "time" when we reduce k). Now consider the model in which we are going forwards time when we increase k, i.e. the model with:

$$
L(AL^{-1}L^{-1} + BL^{-1} + C)d = 0,
$$

subject to the terminal condition that $d_k \to 0$ as $k \to \infty$, which must hold as we have already proved that the first row of M converges to zero. Now, let $z \in \mathbb{C}$, $z \neq 0$ be a solution to:

$$
0 = \det(Az^2 + Bz + C),
$$

and define $\tilde{z} = z^{-1}$, so:

$$\begin{aligned}0 = \det(A + B\tilde{z} + C\tilde{z}^2) &= z^{-2} \det(Az^2 + Bz + C) \\ &= \det(I - G\tilde{z}) \det(B + CF) \det(I\tilde{z} - F).\end{aligned}$$

By Assumption 1, all of the roots of $\det(I\tilde{z} - F)$ are inside the unit circle, thus they cannot contribute to the dynamics of the time reversed process, else the terminal condition would be violated. Thus, the time reversed model has a unique solution satisfying the terminal condition with a transition matrix with the same eigenvalues as G. Consequently, this solution can be calculated via standard methods for solving linear DSGE models, and it will be given by $d_k = Hd_{k-1}$, for all $k > 0$, where $H = -(B + AH)^{-1}C$, and $\phi_H = \phi_G < 1$, by Assumption 2.

It just remains to determine the value of d_0. By the previous results, this must satisfy:

$$-I_{\cdot,1} = Ad_1 + Bd_0 + Cd_{-1} = (AH + B + CF)d_0.$$

Hence:

$$d_0 = -(AH + B + CF)^{-1} I_{\cdot,1}.$$

This gives a readily computed solution for the limits of the diagonals of M. Lastly, note that:

$$|d_{-t,1}| \le \|d_{-t}\|_2 = \|F^t d_0\|_2 \le \|F^t\|_2 \|d_0\|_2 \le \mathscr{C}_{F,\phi_F} \phi_F^t \|d_0\|_2,$$

and:

$$|d_{t,1}| \le \|d_t\|_2 = \|H^t d_0\|_2 \le \|H^t\|_2 \|d_0\|_2 \le \mathscr{C}_{H,\phi_H} \phi_H^t \|d_0\|_2.$$

We will use these results in producing our bounds on ς.

First, fix $T \in \mathbb{N}^+$, and define a new matrix $\underline{M}^{(T)} \in \mathbb{R}^{\mathbb{N}^+ \times \mathbb{N}^+}$ by $\underline{M}^{(T)}_{1:T,1:T} = M_{1:T,1:T}$, and for all $t, k \in \mathbb{N}^+$, with $\min\{t, k\} > T$, $\underline{M}^{(T)}_{t,k} = d_{k-t,1} - \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G}$, then:

$$\begin{aligned}
\varsigma &\ge \max_{\substack{y\in[0,1]^T \\ y_\infty\in[0,1]}} \inf_{t\in\mathbb{N}^+} M_{t,1:\infty} \begin{bmatrix} y \\ y_\infty 1_{\infty\times 1}\end{bmatrix} \ge \max_{\substack{y\in[0,1]^T \\ y_\infty\in[0,1]}} \inf_{t\in\mathbb{N}^+} \underline{M}^{(T)}_{t,1:\infty} \begin{bmatrix} y \\ y_\infty 1_{\infty\times 1}\end{bmatrix} \\
&= \max_{\substack{y\in[0,1]^T \\ y_\infty\in[0,1]}} \min \left\{ \begin{array}{c} \min\limits_{t=1,\dots,T} \left[M_{t,1:T} y + \sum\limits_{k=T+1}^{\infty} \left(d_{k-t,1} - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_\infty \right], \\ \inf\limits_{t\in\mathbb{N}^+, t>T} \left[\sum\limits_{k=1}^{T} \left(d_{k-t,1} - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_k + \sum\limits_{k=T+1}^{\infty} \left(d_{k-t,1} - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_\infty \right] \end{array} \right\} \\
&\ge \max_{\substack{y\in[0,1]^T \\ y_\infty\in[0,1]}} \min \left\{ \begin{array}{c} \min\limits_{t=1,\dots,T} \left[M_{t,1:T} y + ((I-H)^{-1} d_{T+1-t})_1 y_\infty - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} y_\infty \right], \\ \min\limits_{t=T+1,\dots,2T} \left[\begin{array}{c} \sum\limits_{k=1}^{T} \left(d_{-(t-k),1} - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_k + \left((I-F)^{-1} (d_{-1} - d_{-(t-T)}) \right)_1 y_\infty \\ +((I-H)^{-1} d_0)_1 y_\infty - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} y_\infty \end{array} \right], \\ \inf\limits_{t\in\mathbb{N}^+, t>2T} \left[\begin{array}{c} \sum\limits_{k=1}^{T} d_{-(t-k),1} y_k + \left((I-F)^{-1} (d_{-1} - d_{-(t-T)}) \right)_1 y_\infty \\ +((I-H)^{-1} d_0)_1 y_\infty - \mathscr{D}_{\phi_F,\phi_G} \dfrac{\phi_F^{2T+1} \phi_G}{(1-\phi_F\phi_G)(1-\phi_G)} \end{array} \right] \end{array} \right\}.
\end{aligned}$$

Now, for $t \ge T$:

$$\begin{aligned}\left| \left((I-F)^{-1} d_{-(t-T)} \right)_1 \right| &\le \|(I-F)^{-1} d_{-(t-T)}\|_2 \le \|(I-F)^{-1}\|_2 \|d_{-(t-T)}\|_2 \\ &\le \mathscr{C}_{F,\phi_F} \phi_F^{t-T} \|(I-F)^{-1}\|_2 \|d_0\|_2,\end{aligned}$$

so:

$$\begin{aligned}
&\sum_{k=1}^{T} d_{-(t-k),1} y_k - \big((I-F)^{-1} d_{-(t-T)}\big)_1 y_\infty \\
&\qquad\qquad \geq -\sum_{k=1}^{T} \mathscr{C}_{F,\phi_F} \phi_F^{t-k} \|d_0\|_2 - \mathscr{C}_{F,\phi_F} \phi_F^{t-T} \|(I-F)^{-1}\|_2 \|d_0\|_2 y_\infty \\
&\qquad\qquad = -\mathscr{C}_{F,\phi_F} \frac{\phi_F^t(\phi_F^{-T}-1)}{1-\phi_F} \|d_0\|_2 - \mathscr{C}_{F,\phi_F} \phi_F^{t-T} \|(I-F)^{-1}\|_2 \|d_0\|_2 y_\infty,
\end{aligned}$$

thus $\varsigma \geq \underline{\varsigma}$, where:

$$\underline{\varsigma}_T \coloneqq \max_{\substack{y\in[0,1]^T \\ y_\infty \in [0,1]}} \min \left\{ \begin{gathered}
\min_{t=1,\dots,T} \left[M_{t,1:T} y + ((I-H)^{-1} d_{T+1-t})_1 y_\infty - \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} y_\infty \right], \\
\min_{t=T+1,\dots,2T} \left[\begin{gathered} \sum_{k=1}^{T} \left(d_{-(t-k),1} - \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_k + \Big((I-F)^{-1} (d_{-1} - d_{-(t-T)}) \Big)_1 y_\infty \\ + ((I-H)^{-1} d_0)_1 y_\infty - \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} y_\infty \end{gathered} \right], \\
\left[\begin{gathered} -\mathscr{C}_{F,\phi_F} \frac{\phi_F^{2T+1}(\phi_F^{-T}-1)}{1-\phi_F} \|d_0\|_2 - \mathscr{C}_{F,\phi_F} \phi_F^{T+1} \|(I-F)^{-1}\|_2 \|d_0\|_2 y_\infty + ((I-F)^{-1} d_{-1})_1 y_\infty \\ + ((I-H)^{-1} d_0)_1 y_\infty - \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^{2T+1} \phi_G}{(1-\phi_F\phi_G)(1-\phi_G)} \end{gathered} \right]
\end{gathered} \right\}.$$

It is worth noting that as $T \to \infty$, the final minimand in this expression tends to:

$$\big((I-F)^{-1} d_{-1}\big)_1 y_\infty + \big((I-H)^{-1} d_0\big)_1 y_\infty,$$

i.e. a positive multiple of the doubly infinite sum of $d_{1,k}$ over all $k \in \mathbb{Z}$. If this expression is negative, then our lower bound on ς will be negative as well, and hence uninformative.

To construct an upper bound on ς, fix $T \in \mathbb{N}^+$, and define a new matrix $\overline{M}^{(T)} \in \mathbb{R}^{\mathbb{N}^+ \times \mathbb{N}^+}$ by $\overline{M}^{(T)}_{1:T,1:T} = M_{1:T,1:T}$, and for all $t,k \in \mathbb{N}^+$, with $\min\{t,k\} > T$, $\overline{M}^{(T)}_{t,k} = |d_{k-t,1}| + \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G}$. Then:

$$\begin{aligned}
\varsigma &= \sup_{y\in[0,1]^{\mathbb{N}^+}} \inf_{t\in\mathbb{N}^+} M_{t,1:\infty} y \leq \sup_{y\in[0,1]^{\mathbb{N}^+}} \inf_{t\in\mathbb{N}^+} \overline{M}_{t,1:\infty} y \leq \sup_{y\in[0,1]^{\mathbb{N}^+}} \min_{t=1,\dots,T} \overline{M}_{t,1:\infty} y \\
&\leq \max_{y\in[0,1]^T} \min_{t=1,\dots,T} \overline{M}_{t,1:\infty} \begin{bmatrix} y \\ 1_{\infty\times 1} \end{bmatrix} \\
&\leq \max_{y\in[0,1]^T} \min_{t=1,\dots,T} \left[M_{t,1:T} y + \sum_{k=T+1}^{\infty} |d_{k-t,1}| + \sum_{k=T+1}^{\infty} \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right] \\
&\leq \max_{y\in[0,1]^T} \min_{t=1,\dots,T} \left[M_{t,1:T} y + \sum_{k=T+1-t}^{\infty} |d_{k,1}| + \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{1-\phi_F\phi_G} \sum_{k=0}^{\infty} \phi_G^k \right] \\
&\leq \max_{y\in[0,1]^T} \min_{t=1,\dots,T} \left[M_{t,1:T} y + \mathscr{C}_{H,\phi_H} \|d_0\|_2 \phi_H^{T+1-t} \sum_{k=0}^{\infty} \phi_H^k + \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} \right] \\
&= \overline{\varsigma}_T \coloneqq \max_{y\in[0,1]^T} \min_{t=1,\dots,T} \left[M_{t,1:T} y + \frac{\mathscr{C}_{H,\phi_H} \|d_0\|_2 \phi_H^{T+1-t}}{1-\phi_H} + \mathscr{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} \right].
\end{aligned}$$

C. Other properties of the solution set

First, let us give one further definition:

Definition 9 ((Non-)Degenerate matrix) A matrix $M \in \mathbb{R}^{T\times T}$ is called a **non-degenerate matrix** if the principal minors of M are all non-zero. M is called a **degenerate matrix** if it is not a non-degenerate matrix.

Then, conditions for having a finite or convex set of solutions are given in the following propositions.

Proposition 15 The LCP (q, M) has a finite (possibly zero) number of solutions for all $q \in \mathbb{R}^T$ if and only if M is non-degenerate. (Cottle, Pang, and Stone 2009a)

Proposition 16 The LCP (q, M) has a convex (possibly empty) set of solutions for all $q \in \mathbb{R}^T$ if and only if M is column sufficient. (Cottle, Pang, and Stone 2009a)

D. Proof of the sufficient conditions for the existence of a unique solution to the dynamic programming problem

Results when $\tilde{X}$ is possibly non-compact, but $\widetilde{\boldsymbol{\Gamma}}(\boldsymbol{x})$ is compact valued and $\boldsymbol{x} \in \widetilde{\boldsymbol{\Gamma}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \tilde{X}$ We first note that for all $x, z \in \tilde{X}$:

$$\widetilde{\mathscr{F}}(x, z) \le u^{(0)} - \frac{1}{2} u^{(1)} \tilde{u}^{(2)^{-1}} u^{(1)\prime},$$

thus our objective function is bounded above without additional assumptions. For a lower bound, we assume that for all $x \in \tilde{X}$, $x \in \widetilde{\Gamma}(x)$, so holding the state fixed is always feasible. This is true in very many standard applications. Then, the value of setting $x_t = x_0$ for all $t \in \mathbb{N}^+$ provides a lower bound for our objective function.

More precisely, we define $\mathbb{V} := \{v | v: \tilde{X} \to [-\infty, \infty)\}$ and $\underline{v}, \overline{v} \in \mathbb{V}$ by:

$$\underline{v}(x) = \frac{1}{1-\beta} \widetilde{\mathscr{F}}(x_0, x_0),$$

$$\overline{v}(x) = \frac{1}{1-\beta} \left[u^{(0)} - \frac{1}{2} u^{(1)} \tilde{u}^{(2)^{-1}} u^{(1)\prime} \right],$$

for all $x \in \tilde{X}$.

Finally, define $\mathscr{B}: \mathbb{V} \to \mathbb{V}$ by:

$$\mathscr{B}(v)(x) = \sup_{z \in \widetilde{\Gamma}(x)} \left[\widetilde{\mathscr{F}}(x, z) + \beta v(z) \right] \tag{17}$$

for all $v \in \mathbb{V}$ and for all $x \in \tilde{X}$. Then $\mathscr{B}(\underline{v}) \ge \underline{v}$ and $\mathscr{B}(\overline{v}) \le \overline{v}$. Furthermore, if some sequence $(x_t)_{t=1}^{\infty}$ satisfies the constraint that for all $t \in \mathbb{N}^+$, $x_t \in \widetilde{\Gamma}(x_{t-1})$, and the objective in (8) is finite for that sequence, then it must be the case that $\|x_t\|_\infty t \beta^{\frac{t}{2}} \to 0$ as $t \to \infty$ (by the comparison test), so:

$$\liminf_{t\to\infty} \beta^t \underline{v}(x_t) = 0.$$

Additionally, for any sequence $(x_t)_{t=1}^{\infty}$:

$$\limsup_{t\to\infty} \beta^t \overline{v}(x_t) = 0.$$

Thus, our dynamic programming problem satisfies the assumptions of Theorem 2.1 of Kamihigashi (2014), and so $\mathscr{B}$ has a unique fixed point in $[\underline{v}, \overline{v}]$ to which $\mathscr{B}^k(\underline{v})$ converges pointwise, monotonically, as $k \to \infty$, and which is equal to the function $v^*: \tilde{X} \to \mathbb{R}$ defined by:

$$v^*(x_0) = \sup\left\{\sum_{t=1}^{\infty} \beta^{t-1} \widetilde{\mathscr{F}}(x_{t-1}, x_t) \,\middle|\, \forall t \in \mathbb{N}^+, x_t \in \Gamma(x_{t-1})\right\}, \tag{18}$$

for all $x_0 \in \tilde{X}$.

Furthermore, if we define $\mathbb{W} := \{v \in V | v \text{ is continuous on } \tilde{X}, v \text{ is concave on } \tilde{X}\}$, then as $\tilde{u}^{(2)}$ is negative-definite, $\underline{v} \in \mathbb{W}$. Additionally, under the assumption that $\widetilde{\Gamma}(x)$ is compact valued, if $v \in \mathbb{W}$, then $\mathscr{B}(v) \in \mathbb{W}$, by the theorem of the maximum,[23] and, furthermore, there is a unique policy function which attains the supremum in the definition of $\mathscr{B}(v)$. Moreover, $v^* = \lim_{k\to\infty} \mathscr{B}^k(\underline{v})$ is concave and lower semi-continuous on $\tilde{X}$.[24] We just need to prove that v^* is upper semi-continuous.[25] Suppose for a contradiction that it is not, so there exists $x^* \in \tilde{X}$ such that:

$$\limsup_{x\to x^*} v^*(x) > \lim_{k\to\infty} v^*(x^*).$$

Then, there exists $\delta > 0$ such that for all $\epsilon > 0$, there exists $x_0^{(\epsilon)} \in \tilde{X}$ with $\|x^* - x_0^{(\epsilon)}\|_\infty < \epsilon$ such that:

$$v^*\left(x_0^{(\epsilon)}\right) > \delta + v^*(x^*).$$

Now, by the definition of a supremum, for all $\epsilon > 0$, there exists $\left(x_t^{(\epsilon)}\right)_{t=1}^{\infty}$ such that for all $t \in \mathbb{N}^+$, $x_t^{(\epsilon)} \in \Gamma\left(x_{t-1}^{(\epsilon)}\right)$ and:

$$v^*\left(x_0^{(\epsilon)}\right) < \delta + \sum_{t=1}^{\infty} \beta^{t-1} \widetilde{\mathscr{F}}\left(x_{t-1}^{(\epsilon)}, x_t^{(\epsilon)}\right).$$

Hence:

$$\sum_{t=1}^{\infty} \beta^{t-1} \widetilde{\mathscr{F}}\left(x_{t-1}^{(\epsilon)}, x_t^{(\epsilon)}\right) > v^*\left(x_0^{(\epsilon)}\right) - \delta > v^*(x^*).$$

Now, let $\mathscr{S}_0 := \{x \in \tilde{X} | \|x^* - x\|_\infty \le 1\}$, and for $t \in \mathbb{N}^+$, let $\mathscr{S}_t := \Gamma(\mathscr{S}_{t-1})$. Then, since we are assuming Γ is compact valued, for all $t \in \mathbb{N}$, $\mathscr{S}_t$ is compact by the continuity of Γ. Furthermore, for all $t \in \mathbb{N}$ and $\epsilon \in (0,1)$, $x_t^{(\epsilon)} \in \mathscr{S}_t$. Hence, $\prod_{t=0}^{\infty} \mathscr{S}_t$ is sequentially compact in the product topology. Thus, there exists a sequence $(\epsilon_k)_{k=1}^{\infty}$ with $\epsilon_k \to 0$ as $k \to \infty$ and such that $x_t^{(\epsilon_k)}$ converges for all $t \in \mathbb{N}$. Let $x_t := \lim_{k\to\infty} x_t^{(\epsilon_k)}$,

[23] See e.g. Theorem 3.6 and following of Stokey, Lucas, and Prescott (1989).

[24] See e.g. Lemma 2.41 of Aliprantis and Border (2013).

[25] In the following, we broadly follow the proof of Lemma 3.3 of Kamihigashi and Roy (2003).

and note that $x^* = x_0 \in \mathscr{S}_0 \subseteq \tilde{X}$, and that for all $t, k \in \mathbb{N}^+$, $x_t^{(\epsilon_k)} \in \Gamma\left(x_{t-1}^{(\epsilon_k)}\right)$, so by the continuity of Γ, $x_t \in \Gamma(x_{t-1})$ for all $t \in \mathbb{N}^+$. Thus, by Fatou's Lemma:

$$v^*(x^*) \geq \sum_{t=1}^{\infty} \beta^{t-1} \widetilde{\mathscr{F}}(x_{t-1}, x_t) \geq \limsup_{k\to\infty} \sum_{t=1}^{\infty} \beta^{t-1} \widetilde{\mathscr{F}}\left(x_{t-1}^{(\epsilon,k)}, x_t^{(\epsilon,k)}\right) > v^*(x^*),$$

which gives the required contradiction. Thus v^* is continuous and concave, and there is a unique policy function which attains the supremum in the definition of $\mathscr{B}(v^*) = v^*$.

Results when $\tilde{X}$ is compact If $\tilde{X}$ is compact, then Γ is compact valued. Furthermore, $\tilde{X}$ is clearly convex, and Γ is continuous. Thus assumption 4.3 of Stokey, Lucas, and Prescott (1989) (henceforth: SLP) is satisfied. Since the continuous image of a compact set is compact, $\widetilde{\mathscr{F}}$ is bounded above and below, so assumption 4.4 of SLP is satisfied as well. Furthermore, $\widetilde{\mathscr{F}}$ is concave and Γ is convex, so assumptions 4.7 and 4.8 of SLP are satisfied too. Thus, by theorem 4.6 of SLP, with $\mathscr{B}$ defined as in equation (17) and v^* defined as in equation (18), $\mathscr{B}$ has a unique fixed point which is continuous and equal to v^*. Moreover, by theorem 4.8 of SLP, there is a unique policy function which attains the supremum in the definition of $\mathscr{B}(v^*) = v^*$.

E. Proof of the sufficiency of the KKT and limit conditions

Suppose that $(x_t)_{t=1}^{\infty}$, $(\lambda_t)_{t=1}^{\infty}$ satisfy the KKT conditions given in equations (10) and (11), and that $x_t \to \mu$ and $\lambda_t \to \overline{\lambda}$ as $t \to \infty$. Let $(z_t)_{t=0}^{\infty}$ satisfy $z_0 = x_0$ and $z_t \in \widetilde{\Gamma}(z_{t-1})$ for all $t \in \mathbb{N}^+$. Then, by the KKT conditions and the concavity of:

$$(x_{t-1}, x_t) \mapsto \widetilde{\mathscr{F}}(x_{t-1}, x_t) + \lambda_t' \left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}\right],$$

we have that for all $T \in \mathbb{N}^+$:[26]

$$\begin{aligned}
&\sum_{t=1}^{T} \beta^{t-1}\left[\widetilde{\mathscr{F}}(x_{t-1}, x_t) - \widetilde{\mathscr{F}}(z_{t-1}, z_t)\right] \\
&= \sum_{t=1}^{T} \beta^{t-1}\left[\widetilde{\mathscr{F}}(x_{t-1}, x_t) + \lambda_t'\left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}\right] - \widetilde{\mathscr{F}}(z_{t-1}, z_t)\right] \\
&\geq \sum_{t=1}^{T} \beta^{t-1}\left[\widetilde{\mathscr{F}}(x_{t-1}, x_t) + \lambda_t'\left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}\right] - \widetilde{\mathscr{F}}(z_{t-1}, z_t)\right. \\
&\qquad\qquad \left. - \lambda_t'\left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} z_{t-1} - \mu \\ z_t - \mu \end{bmatrix}\right]\right] \\
&\geq \sum_{t=1}^{T} \beta^{t-1}\left[\left[u_{\cdot,2}^{(1)} + \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}' \tilde{u}_{\cdot,2}^{(2)} + \lambda_t' \Psi_{\cdot,2}^{(1)}\right](x_t - z_t)\right. \\
&\qquad\qquad \left. + \left[u_{\cdot,1}^{(1)} + \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}' \tilde{u}_{\cdot,1}^{(2)} + \lambda_t' \Psi_{\cdot,1}^{(1)}\right](x_{t-1} - z_{t-1})\right]
\end{aligned}$$

[26] Here, we broadly follow the proof of Theorem 4.15 of Stokey, Lucas, and Prescott (1989).

$$
\begin{aligned}
&= \sum_{t=1}^{T} \beta^{t-1} \Bigg[\Bigg[u_{\cdot,2}^{(1)} + \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}' \tilde{u}_{\cdot,2}^{(2)} + \lambda_t' \Psi_{\cdot,2}^{(1)} \\
&\qquad\qquad + \beta \left[u_{\cdot,1}^{(1)} + \begin{bmatrix} x_t - \mu \\ x_{t+1} - \mu \end{bmatrix}' \tilde{u}_{\cdot,1}^{(2)} + \lambda_{t+1}' \Psi_{\cdot,1}^{(1)} \right] \Bigg] (x_t - z_t) \Bigg] \\
&\qquad\qquad + \beta^T \left[u_{\cdot,1}^{(1)} + \begin{bmatrix} x_T - \mu \\ x_{T+1} - \mu \end{bmatrix}' \tilde{u}_{\cdot,1}^{(2)} + \lambda_{T+1}' \Psi_{\cdot,1}^{(1)} \right] (z_T - x_T) \\
&= \beta^T \left[u_{\cdot,1}^{(1)} + \begin{bmatrix} x_T - \mu \\ x_{T+1} - \mu \end{bmatrix}' \tilde{u}_{\cdot,1}^{(2)} + \lambda_{T+1}' \Psi_{\cdot,1}^{(1)} \right] (z_T - x_T).
\end{aligned}
$$

Thus:

$$
\begin{aligned}
&\sum_{t=1}^{\infty} \beta^{t-1} \big[\tilde{\mathcal{F}}(x_{t-1}, x_t) - \tilde{\mathcal{F}}(z_{t-1}, z_t) \big] \\
&\qquad \geq \lim_{T\to\infty} \beta^T \left[u_{\cdot,1}^{(1)} + \begin{bmatrix} x_T - \mu \\ x_{T+1} - \mu \end{bmatrix}' \tilde{u}_{\cdot,1}^{(2)} + \lambda_{T+1}' \Psi_{\cdot,1}^{(1)} \right] (z_T - x_T) \\
&\qquad = \lim_{T\to\infty} \beta^T \big[u_{\cdot,1}^{(1)} + \overline{\lambda}' \Psi_{\cdot,1}^{(1)} \big] (z_T - \mu) = \lim_{T\to\infty} \beta^T \big[u_{\cdot,1}^{(1)} + \overline{\lambda}' \Psi_{\cdot,1}^{(1)} \big] z_T.
\end{aligned}
$$

Now, suppose $\lim_{T\to\infty} \beta^T z_T \neq 0$, then since $\tilde{u}^{(2)}$ is negative definite:

$$
\sum_{t=1}^{\infty} \beta^{t-1} \tilde{\mathcal{F}}(z_{t-1}, z_t) = -\infty,
$$

so $(z_t)_{t=0}^{\infty}$ cannot be optimal. Hence, regardless of the value of $\lim_{T\to\infty} \beta^T z_T$:

$$
\sum_{t=1}^{\infty} \beta^{t-1} \big[\tilde{\mathcal{F}}(x_{t-1}, x_t) - \tilde{\mathcal{F}}(z_{t-1}, z_t) \big] \geq 0,
$$

which implies that $(x_t)_{t=1}^{\infty}$ solves Problem 5.

F. Results from and for general dynamic programming problems

Here we consider non-linear dynamic programming problems with general objective functions. Consider then the following generalisation of Problem 5:

Problem 6 Suppose $\Gamma: \mathbb{R}^n \to \mathbb{P}(\mathbb{R}^n)$ is a given compact, convex valued continuous function. Define $X \coloneqq \{x \in \mathbb{R}^n | \Gamma(x) \neq \emptyset\}$, and suppose without loss of generality that for all $x \in \mathbb{R}^n$, $\Gamma(x) \cap X = \Gamma(x)$. Further suppose that $\mathcal{F}: X \times X \to \mathbb{R}$ is a given twice continuously differentiable, concave function, and that $x_0 \in X$ and $\beta \in (0,1)$ are given. Choose $x_1, x_2, \ldots$ to maximise:

$$
\liminf_{T\to\infty} \sum_{t=1}^{T} \beta^{t-1} \mathcal{F}(x_{t-1}, x_t),
$$

subject to the constraints that for all $t \in \mathbb{N}^+$, $x_t \in \Gamma(x_{t-1})$.

For tractability, we make the following additional assumption, which enables us to uniformly approximate Γ by a finite number of inequalities:

Assumption 4 X is compact.

Then, by theorem 4.8 of Stokey, Lucas, and Prescott (1989), there is a unique solution to Problem 6 for any x_0. We further assume the following to ensure that there is a natural point to approximate around:[27]

Assumption 5 There exists $\mu \in X$ such that for any given $x_0 \in X$, in the solution to Problem 6 with that x_0, as $t \to \infty$, $x_t \to \mu$.

Having defined μ, we can let $\tilde{\mathscr{F}}$ be a second order Taylor approximation to $\mathscr{F}$ around μ, which will take the form of equation (7). Assumption 3 will be satisfied for this approximation thanks to the concavity of $\mathscr{F}$. To apply the previous results, we also then need to approximate the constraints.

Suppose first that the graph of Γ is convex, i.e. the set $\{(x, z) | x \in X, z \in \Gamma(x)\}$ is convex. Since it is also compact, by Assumption 4, for any $\epsilon > 0$, there exists $c \in \mathbb{N}$, $\Psi^{(0)} \in \mathbb{R}^{c \times 1}$ and $\Psi^{(1)} \in \mathbb{R}^{c \times 2n}$ such that with $\tilde{\Gamma}$ defined as in equation (5) and $\tilde{X}$ defined as in equation (6):

1) $\mu \in \tilde{X} \subseteq X$,
2) for all $x \in X$, there exists $\tilde{x} \in \tilde{X}$ such that $\|x - \tilde{x}\|_2 < \epsilon$,
3) for all $x \in \tilde{X}$, $\tilde{\Gamma}(x) \subseteq \Gamma(x)$,
4) for all $x \in \tilde{X}$, and for all $z \in \Gamma(x)$, there exists $\tilde{z} \in \tilde{\Gamma}(x)$ such that $\|z - \tilde{z}\|_2 < \epsilon$.

(This follows from standard properties of convex sets.) Then, by our previous results, the following proposition is immediate:

Proposition 17 Suppose we are given a problem in the form of Problem 6 (and which satisfies Assumption 4 and Assumption 5). If the graph of Γ is convex, then we can construct a problem in the form of the multiple-bound generalisation of Problem 2 which encodes a local approximation to the original dynamic programming problem around $x_t = \mu$. Furthermore, the LCP corresponding to this approximation will have a unique solution for all $x_0 \in \tilde{X}$. Moreover, the approximation is consistent for quadratic objectives in the sense that as the number of inequalities used to approximate Γ goes to infinity, the approximate value function converges uniformly to the true value function.

Unfortunately, if the graph of Γ is non-convex, then we will not be able to derive similar results. To see the best we could do along similar proof lines, here we merely sketch the construction of an approximation to the graph of Γ in this case. We will need to assume that there exists $z \in \operatorname{int} \Gamma(x)$ for all $x \in X$, which precludes the existence of equality constraints.[28] We first approximate the graph of Γ by a polytope (i.e. n

[27] If X is convex, then the existence of a fixed point of the policy function is a consequence of the Brouwer fixed point theorem, but there is no reason the fixed point guaranteed by Brouwer's theorem should be even locally attractive.
[28] This is often not too much of a restriction, since equality constraints may be substituted out.

dimensional polygon) contained in the graph of Γ such that all points in the graph of Γ are within $\frac{\epsilon}{2}$ of a point in the polytope. Then, providing ϵ is sufficiently small, for each simplicial surface element of the polytope, indexed by $k \in \{1, \dots, c\}$, we can find a quadratic function $q_k : X \times X \to \mathbb{R}$ with:

$$q_k = \Psi_k^{(0)} + \Psi_{k,\cdot}^{(1)} \begin{bmatrix} x - \mu \\ z - \mu \end{bmatrix} + \begin{bmatrix} x - \mu \\ z - \mu \end{bmatrix}' \Psi_k^{(2)} \begin{bmatrix} x - \mu \\ z - \mu \end{bmatrix}$$

for all $x, z \in X$ and such that q_k is zero at the corners of the simplicial surface element, such that q_k is weakly negative on its surface, such that $\Psi_k^{(2)}$ is symmetric positive definite, and such that all points in the polytope are within $\frac{\epsilon}{2}$ of a point in the set:

$$\{(x, z) \in X \times X | \forall k \in \{1, \dots, S\}, 0 \le q_k(x, z)\}.$$

This gives a set of quadratic constraints that approximate Γ. If we then define:

$$\tilde{u}^{(2)} := u^{(2)} + \sum_{k=1}^{c} \overline{\lambda}'_{\Psi,k} \Psi_k^{(2)},$$

where $u^{(2)}$ is the Hessian of $\mathscr{F}$, then the Lagrangian in equation (9) is the same as what would be obtained from taking a second order Taylor approximation to the Lagrangian of the problem of maximising our non-linear objective subject to the approximate quadratic constraints, suggesting it may perform acceptably well for x near μ, along similar lines to the results of Levine, Pearlman, and Pierse (2008) and Benigno and Woodford (2012). However, existence of a unique solution to the original problem cannot be used to establish even the existence of a solution of the approximated problem, since only linear approximations to the quadratic constraints would be imposed by our algorithm, giving a greatly reduced choice set (as the quadratic terms are positive definite).

G. Proof of the properties of the BPY model

Defining $x_t = [x_{i,t} \quad x_{y,t} \quad x_{\pi,t}]'$, the BPY model is in the form of Problem 2, with:

$$A := \begin{bmatrix} 0 & -\alpha_{\Delta y} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B := \begin{bmatrix} -1 & \alpha_{\Delta y} & \alpha_\pi \\ -\frac{1}{\sigma} & -1 & 0 \\ 0 & \gamma & -1 \end{bmatrix}, \qquad C := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\sigma} \\ 0 & 0 & \beta \end{bmatrix}.$$

Assumption 2 is satisfied for this model as:

$$\det(A + B + C) = \det \begin{bmatrix} -1 & 0 & \alpha_\pi \\ -\frac{1}{\sigma} & 0 & \frac{1}{\sigma} \\ 0 & \gamma & -1 \end{bmatrix} \neq 0$$

as $\alpha_\pi \neq 1$ and $\gamma \neq 0$. Let $f := F_{2,2}$, where F is as in Assumption 1. Then:

$$F = \begin{bmatrix} 0 & \alpha_{\Delta y}(f - 1) + \alpha_\pi \frac{\gamma f}{1 - \beta f} & 0 \\ 0 & f & 0 \\ 0 & \frac{\gamma f}{1 - \beta f} & 0 \end{bmatrix}.$$

Hence:

$$f = f^2 - \frac{1}{\sigma}\left(\alpha_{\Delta y}(f-1) + \alpha_\pi \frac{\gamma f}{1-\beta f} - \frac{\gamma f^2}{1-\beta f}\right),$$

i.e.:

$$\beta\sigma f^3 - \left(\left(\alpha_{\Delta y} + \sigma\right)\beta + \gamma + \sigma\right)f^2 + \left((1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma\right)f - \alpha_{\Delta y} = 0. \quad (19)$$

When $f \le 0$, the left hand side is negative, and when $f = 1$, the left hand side equals $(\alpha_\pi - 1)\gamma > 0$ (by assumption on α_π), hence equation (3) has either one or three solutions in $(0,1)$, and no solutions in $(-\infty, 0]$. We wish to prove there is a unique solution in $(-1,1)$. First note that when $\alpha_\pi = 1$, the discriminant of the polynomial is:

$$\left((1-\beta)\left(\alpha_{\Delta y} - \sigma\right) - \gamma\right)^2 \left(\left(\beta\alpha_{\Delta y}\right)^2 + 2\beta(\gamma - \sigma)\alpha_{\Delta y} + (\gamma + \sigma)^2\right).$$

The first multiplicand is positive. The second is minimised when $\sigma = \beta\alpha_{\Delta y} - \gamma$, at the value $4\beta\gamma\alpha_{\Delta y} > 0$, hence this multiplicand is positive too. Consequently, at least for small α_π, there are three real solutions for f, so there may be multiple solutions in $(0,1)$.

Suppose for a contradiction that there were at least three solutions to equation (3) in $(0,1)$ (double counting repeated roots), even for arbitrary large $\beta \in (0,1)$. Let $f_1, f_2, f_3 \in (0,1)$ be the three roots. Then, by Vieta's formulas:

$$3 > f_1 + f_2 + f_3 = \frac{\left(\alpha_{\Delta y} + \sigma\right)\beta + \gamma + \sigma}{\beta\sigma},$$
$$3 > f_1 f_2 + f_1 f_3 + f_2 f_3 = \frac{(1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma}{\beta\sigma},$$
$$1 > f_1 f_2 f_3 = \frac{\alpha_{\Delta y}}{\beta\sigma},$$

so:

$$(2\beta - 1)\sigma > \beta\alpha_{\Delta y} + \gamma > \gamma > 0$$
$$\beta > \frac{1}{2}, \qquad (2\beta - 1)\sigma > \gamma,$$
$$\beta\sigma > \beta\alpha_{\Delta y} + \gamma + \sigma(1-\beta),$$
$$2\beta\sigma > (1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma(1-\beta),$$
$$\beta\sigma > \alpha_{\Delta y}.$$

Also, the first derivative of equation (3) must be strictly positive at $f = 1$, so:

$$(1-\beta)\left(\alpha_{\Delta y} - \sigma\right) + (\alpha_\pi - 2)\gamma > 0.$$

Combining all of these inequalities gives the bounds:

$$0 < \alpha_{\Delta y} < 2\sigma - \frac{\gamma + \sigma}{\beta},$$
$$2 + \frac{(1-\beta)\left(\sigma - \alpha_{\Delta y}\right)}{\gamma} < \alpha_\pi < \frac{(3\beta - 1)\sigma - (1+\beta)\alpha_{\Delta y}}{\gamma}.$$

Furthermore, if there are multiple solutions to equation (3), then the discriminant of its first derivative must be weakly positive, i.e.:

$$\left(\left(\alpha_{\Delta y} + \sigma\right)\beta + \gamma + \sigma\right)^2 - 3\beta\sigma\left((1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma\right) \ge 0.$$

Therefore, we have the following bounds on α_π:

$$2+\frac{(1-\beta)\big(\sigma-\alpha_{\Delta y}\big)}{\gamma}<\alpha_\pi\le\frac{\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)^2-3\beta\sigma\Big((1+\beta)\alpha_{\Delta y}+\sigma\Big)}{3\beta\sigma\gamma}$$

since,

$$\begin{aligned}\frac{(3\beta-1)\sigma-(1+\beta)\alpha_{\Delta y}}{\gamma}&-\frac{\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)^2-3\beta\sigma\Big((1+\beta)\alpha_{\Delta y}+\sigma\Big)}{3\beta\sigma\gamma}\\&=\frac{\Big(\big(2\sigma-\alpha_{\Delta y}\big)\beta-\gamma-\sigma\Big)\Big(\big(4\sigma+\alpha_{\Delta y}\big)\beta+\gamma+\sigma\Big)}{3\beta\gamma\sigma}>0\end{aligned}$$

as $\alpha_{\Delta y}<2\sigma-\frac{\gamma+\sigma}{\beta}$. Consequently, there exists $\lambda,\mu,\kappa\in[0,1]$ such that:

$$\begin{aligned}\alpha_\pi&=(1-\lambda)\left[2+\frac{(1-\beta)\big(\sigma-\alpha_{\Delta y}\big)}{\gamma}\right]\\&\quad+\lambda\left[\frac{\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)^2-3\beta\sigma\Big((1+\beta)\alpha_{\Delta y}+\sigma\Big)}{3\beta\sigma\gamma}\right],\\ \alpha_{\Delta y}&=(1-\mu)[0]+\mu\left[2\sigma-\frac{\gamma+\sigma}{\beta}\right],\\ \gamma&=(1-\kappa)[0]+\kappa[(2\beta-1)\sigma]\end{aligned}$$

These simultaneous equations have unique solutions for α_π, $\alpha_{\Delta y}$ and γ in terms of λ, μ and κ. Substituting these solutions into the discriminant of equation (3) gives a polynomial in $\lambda,\mu,\kappa,\beta,\sigma$. As such, an exact global maximum of the discriminant may be found subject to the constraints $\lambda,\mu,\kappa\in[0,1]$, $\beta\in\left[\frac{1}{2},1\right]$, $\sigma\in[0,\infty)$, by using an exact compact polynomial optimisation solver, such as that in the Maple computer algebra package. Doing this gives a maximum of 0 when $\beta\in\left\{\frac{1}{2},1\right\}$, $\kappa=1$ and $\sigma=0$. But of course, we actually require that $\beta\in\left(\frac{1}{2},1\right)$, $\kappa<1$, $\sigma>0$. Thus, by continuity, the discriminant is strictly negative over the entire possible domain. This gives the required contradiction to our assumption of three roots to the polynomial, establishing that Assumption 1 holds for this model.

Now, when $T=1$, M is equal to the top left element of the matrix $-(B+CF)^{-1}$, i.e.:

$$M=\frac{\beta\sigma f^2-\big((1+\beta)\sigma+\gamma\big)f+\sigma}{\beta\sigma f^2-\Big((1+\beta)\sigma+\gamma+\beta\alpha_{\Delta y}\Big)f+\sigma+\alpha_{\Delta y}+\gamma\alpha_\pi}.$$

Now, multiplying the denominator by f gives:

$$\begin{aligned}\beta\sigma f^3&-\Big((1+\beta)\sigma+\gamma+\beta\alpha_{\Delta y}\Big)f^2+\big(\sigma+\alpha_{\Delta y}+\gamma\alpha_\pi\big)f\\&=\Big[\beta\sigma f^3-\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)f^2+\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma\Big)f\\&\quad-\alpha_{\Delta y}\Big]-\big[\beta\alpha_{\Delta y}f-\alpha_{\Delta y}\big]=(1-\beta f)\alpha_{\Delta y}>0,\end{aligned}$$

by equation (19). Hence, the sign of M is that of $\beta\sigma f^2 - \big((1+\beta)\sigma+\gamma\big)f+\sigma$. I.e., M is negative if and only if:

$$\frac{\big((1+\beta)\sigma+\gamma\big)-\sqrt{\big((1+\beta)\sigma+\gamma\big)^2-4\beta\sigma^2}}{2\beta\sigma}<f$$
$$<\frac{\big((1+\beta)\sigma+\gamma\big)+\sqrt{\big((1+\beta)\sigma+\gamma\big)^2-4\beta\sigma^2}}{2\beta\sigma}.$$

The upper limit is greater than 1, so only the lower is relevant. To translate this bound on f into a bound on $\alpha_{\Delta y}$, we first need to establish that f is monotonic in $\alpha_{\Delta y}$.

Totally differentiating equation (19) gives:

$$\left[3\beta\sigma f^2-2\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)f+\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma\Big)\right]\frac{df}{d\alpha_{\Delta y}}$$
$$=(1-\beta f)(1-f)>0.$$

Thus, the sign of $\frac{df}{d\alpha_{\Delta y}}$ is equal to that of:

$$3\beta\sigma f^2-2\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)f+\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma\Big).$$

Note, however, that this expression is just the derivative of the left hand side of equation (19) with respect to f.

To establish the sign of $\frac{df}{d\alpha_{\Delta y}}$, we consider two cases. First, suppose that equation (19) has three real solutions. Then, the unique solution to equation (19) in $(0,1)$ is its lowest solution. Hence, this solution must be below the first local maximum of the left hand side of equation (19). Consequently, at the $f\in(0,1)$, which solves equation (19), $3\beta\sigma f^2-2\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)f+\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma\Big)>0$. Alternatively, suppose that equation (19) has a unique real solution. Then the left hand side of this equation cannot change sign in between its local maximum and its local minimum (if it has any). Thus, at the $f\in(0,1)$ at which it changes sign, we must have that $3\beta\sigma f^2-2\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)f+\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma\Big)>0$. Therefore, in either case $\frac{df}{d\alpha_{\Delta y}}>0$, meaning that f is monotonic increasing in $\alpha_{\Delta y}$.

Consequently, to find the critical $\big(f,\alpha_{\Delta y}\big)$ at which M changes sign, it is sufficient to find the lowest solution with respect to both f and $\alpha_{\Delta y}$ of the pair of equations:

$$\beta\sigma f^2-\big((1+\beta)\sigma+\gamma\big)f+\sigma=0,$$
$$\beta\sigma f^3-\Big(\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma\Big)f^2+\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma\Big)f-\alpha_{\Delta y}=0.$$

The former implies that:

$$\beta\sigma f^3-\big((1+\beta)\sigma+\gamma\big)f^2+\sigma f=0,$$

so, by the latter:

$$\alpha_{\Delta y}\beta f^2-\Big((1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi\Big)f+\alpha_{\Delta y}=0.$$

If $\alpha_{\Delta y}=\sigma\alpha_\pi$, then this equation holds if and only if:

$$\sigma\beta f^2-\big((1+\beta)\sigma+\gamma\big)f+\sigma=0.$$

Therefore, the critical $\left(f, \alpha_{\Delta y}\right)$ at which M changes sign are given by:

$$\alpha_{\Delta y} = \sigma \alpha_{\pi},$$
$$f = \frac{\left((1+\beta)\sigma + \gamma\right) - \sqrt{\left((1+\beta)\sigma + \gamma\right)^2 - 4\beta\sigma^2}}{2\beta\sigma}.$$

Thus, M is negative if and only if $\alpha_{\Delta y} > \sigma \alpha_{\pi}$, and M is zero if and only if $\alpha_{\Delta y} = \sigma \alpha_{\pi}$.

H. Proof of the properties of the BPY model with level targeting

Defining $x_t = [x_{i,t} \quad x_{y,t} \quad x_{p,t}]'$, the model of section 3.3 is in the form of Problem 2, with:

$$A := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad B := \begin{bmatrix} -1 & \alpha_{\Delta y} & \alpha_{\pi} \\ -\frac{1}{\sigma} & -1 & -\frac{1}{\sigma} \\ 0 & \gamma & -1-\beta \end{bmatrix}, \qquad C := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\sigma} \\ 0 & 0 & \beta \end{bmatrix}.$$

Assumption 2 is satisfied for this model as:

$$\det(A + B + C) = \det \begin{bmatrix} -1 & \alpha_{\Delta y} & \alpha_{\pi} \\ -\frac{1}{\sigma} & 0 & 0 \\ 0 & \gamma & -1 \end{bmatrix} \neq 0$$

as $\alpha_{\Delta y} \neq 0$ and $\alpha_{\pi} \neq 0$. Let $f := F_{3,3}$, where F is as in Assumption 1. Then:

$$F = \begin{bmatrix} 0 & 0 & \frac{f(1-f)\left(\sigma\alpha_{\pi} - \alpha_{\Delta y}\right)}{\alpha_{\Delta y} + (1-f)\sigma} \\ 0 & 0 & \frac{f(1-f-\alpha_{\pi})}{\alpha_{\Delta y} + (1-f)\sigma} \\ 0 & 0 & f \end{bmatrix},$$

and so:

$$\beta\sigma f^3 - \left((1+2\beta)\sigma + \beta\alpha_{\Delta y} + \gamma\right)f^2 + \left((2+\beta)\sigma + (1+\beta)\alpha_{\Delta y} + (1+\alpha_{\pi})\gamma\right)f - \left(\sigma + \alpha_{\Delta y}\right) = 0.$$

Now define:

$$\hat{\alpha}_{\Delta y} := \sigma + \alpha_{\Delta y}, \qquad \hat{\alpha}_{\pi} := 1 + \alpha_{\pi}$$

so:

$$\beta\sigma f^3 - \left(\left(\hat{\alpha}_{\Delta y} + \sigma\right)\beta + \gamma + \sigma\right)f^2 + \left((1+\beta)\hat{\alpha}_{\Delta y} + \gamma\hat{\alpha}_{\pi} + \sigma\right)f - \hat{\alpha}_{\Delta y} = 0.$$

This is identical to the equation for f in the previous section, apart from the fact that $\hat{\alpha}_{\Delta y}$ has replaced $\alpha_{\Delta y}$ and $\hat{\alpha}_{\pi}$ has replaced α_{π}. Hence, by the results of the previous section, Assumption 1 holds for this model as well.

Finally, for this model, with $T = 1$, we have that:

$$M = \frac{(1-f)(1+(1-f)\beta)\sigma^2 + \left((1+(1-f)\beta)\alpha_{\Delta y} + \left((1-f) + \alpha_{\pi}f\right)\gamma\right)\sigma + (1-f)\gamma\alpha_{\Delta y}}{\left((1-f)(1+(1-f)\beta)\sigma + (1+(1-f)\beta)\alpha_{\Delta y} + \left((1-f) + \alpha_{\pi}\right)\gamma\right)\left(\sigma + \alpha_{\Delta y}\right)} > 0.$$