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

Tom D. Holden, *Deutsche Bundesbank*[*]

Contents

---

[*] Address: Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany.

Email: thomas.holden@gmail.com. Website: https://www.tholden.org/. The views expressed in this paper are those of the author and do not represent the views of the Deutsche Bundesbank, the Eurosystem or its staff.

The author would like to thank all those who have commented on the paper and related work, including but not limited to: Gary Anderson, Martin Andreasen, Paul Beaurdry, Saroj Bhattarai, Florin Bilbiie, Charles Brendon, Olivier Coibion, Zsolt Csizmadia, Oliver de Groot, Michael Devereux, Ana-Maria Dumitru, Bill Dupor, Roger Farmer, Michael Funke, William Gavin, Fabio Ghironi, Andy Glover, Seetharama Gowda, Luca Guerrieri, Pablo Guerrón-Quintana, Matteo Iacoviello, Tibor Illés, Michel Juillard, Hong Lan, Paul Levine, Vivien Lewis, Jesper Lindé, Albert Marcet, Enrique Martinez-Garcia, Antonio Mele, Alexander Meyer-Gohde, Virgiliu Midrigan, Boris Mordukhovich, Adrienn Nagy, Jong-Shi Pang, Matthias Paustian, Jesse Perla, Frank Portier, Søren Ravn, Alexander Richter, Chris Sims, Jonathan Swarbrick, Harald Uhlig, Simon Wren-Lewis and Carlos Zarazaga.

The author gratefully acknowledges financial support from the ESRC and the EC, and the provision of office space by the University of Surrey and the University of Washington, in which parts of this paper were written.

The research leading to these results has received funding from the European Commission’s Seventh Framework Programme (FP7/2007-2013) under grant agreement “Integrated Macro-Financial Modelling for Robust Policy Design” (MACFINROBODS, grant no. 612796).

Appendix A: Getting started with DynareOBC

DynareOBC is a MATLAB toolbox designed to simulate and analyse models with occasionally binding constraints. It relies on Dynare (Adjemian et al. 2011) internally. To get started with DynareOBC, first download the latest release (dynareOBCRelease.zip) from:

https://github.com/tholden/dynareOBC/releases

Extract the zip archive into a sub-folder. You should also install the latest stable version of Dynare from:

http://www.dynare.org/download/dynare-stable

While DynareOBC contains a MILP solver, for best results, at this point, you should install a commercial MILP solver. Many of these are free for academics. We have had good results with Gurobi, which is available for academics by following the steps here:

http://www.gurobi.com/academia/for-universities

Other MILP solvers which are available for free to academics are documented in DynareOBC's ReadMe.pdf.

If you do not have administrative rights on your machine, you will also need to get your administrator to install a few minor dependencies for you, which otherwise DynareOBC would install itself. Full instructions for this are given in DynareOBC's ReadMe.pdf.

Next, open MATLAB, reset the MATLAB path (to be on the safe side) and then add only the following folders to your path. In each case, you should not click "add with subfolders". Only the folders specified need adding:

1) The "matlab" folder within Dynare.
2) The root folder of DynareOBC, i.e. the folder containing "dynareOBC.m".
3) The "matlab" folder within whichever MILP solver you installed (if any).

You can now test your set-up of DynareOBC by typing:

dynareOBC TestSolvers

at the MATLAB command prompt. The first time you run DynareOBC it will install various dependencies, and it may restart MATLAB several times. Note that if you have not installed a commercial MILP solver, you should say "yes" when offered the choice to install "SCIP", otherwise DynareOBC's performance will be severely compromised. When DynareOBC has installed everything necessary, it will run the solver tests. Double check in particular that the LP and MILP tests are passing. (Results for the other tests, e.g. semi-definite programming are not relevant.)

If everything has worked up to this point, then you now have a fully functioning install of DynareOBC. To see it in action, you could start by running DynareOBC's included examples. Most of these examples can be run by changing to DynareOBC's "Examples" directory in MATLAB, and then executing the script "RunAllExamples". This iterates over the various sub-directories of the "Examples" directory, running the script "RunExample" within each.

Developing your own models for use with DynareOBC is easy. You can include one or more occasionally binding constraints directly within your MOD file. For example, to include a zero lower bound on nominal interest rates, your MOD file might contain the line:

```
i = max( 0, 1.5 * pi + 0.25 * y );
```

DynareOBC supports both max and min (with two arbitrary arguments) and abs (with one arbitrary argument). There are no restrictions on what is contained within the brackets. You do not have to have a 0 term, and it does not matter which of the arguments of max or min is bigger or smaller in steady state. The only limitation is that the two arguments of max or min cannot be identical in steady state (likewise, the argument of abs cannot be zero in steady state). For a work-around of this limitation in a financial frictions context, see the approach of Swarbrick, Holden & Levine (2016).

Once you have included an OBC in you MOD file, you can run it with DynareOBC by typing:

```
dynareOBC ModFileName.mod
```

where "ModFileName.mod" is the name of your MOD file. Just as with standard Dynare, if you have specified e.g. "irf=40" within your stoch_simul command, then DynareOBC will produce impulse responses. Likewise, if you have specified e.g. "periods=1000" within your stoch_simul command, then DynareOBC will produce a stochastic simulation.

As an example, the file "bbw2016.mod" in the "Examples/BonevaBraunWaki2016" directory of DynareOBC contains the line:

r = max(0, re + phi_pi * (pi - pi_STEADY) + phi_y * (gdp - gdp_STEADY));

in its model block, and has the following stoch_simul command:

stoch_simul(order = 1, periods = 0, irf = 40);

We can run the MOD file (implementing the model of Boneva, Braun & Waki (2016)) by typing:

dynareOBC bbw2016.mod

from within the "Examples/BonevaBraunWaki2016" directory. Doing this produces two sets of impulse responses, however none of them hit the zero lower bound, as the shock is too small. (This uses the solution algorithm of Holden (2016).)

To produce impulse responses to a larger shock, we can run DynareOBC with the ShockScale command line option. This increases the size of the initial impulse in IRF generation, without altering the standard deviations of the model's shocks, or otherwise changing the behaviour of stochastic simulation. For example, if we run:

dynareOBC bbw2016.mod ShockScale=5

then DynareOBC produces IRFs to a 5 standard deviation shock to each of the model's exogenous variables. This produces the two plots shown in Figure 1. In all DynareOBC plots, the solid line shows the economy's path imposing the bound(s), and the dotted line shows the path the economy would have taken were it not for the bound(s).

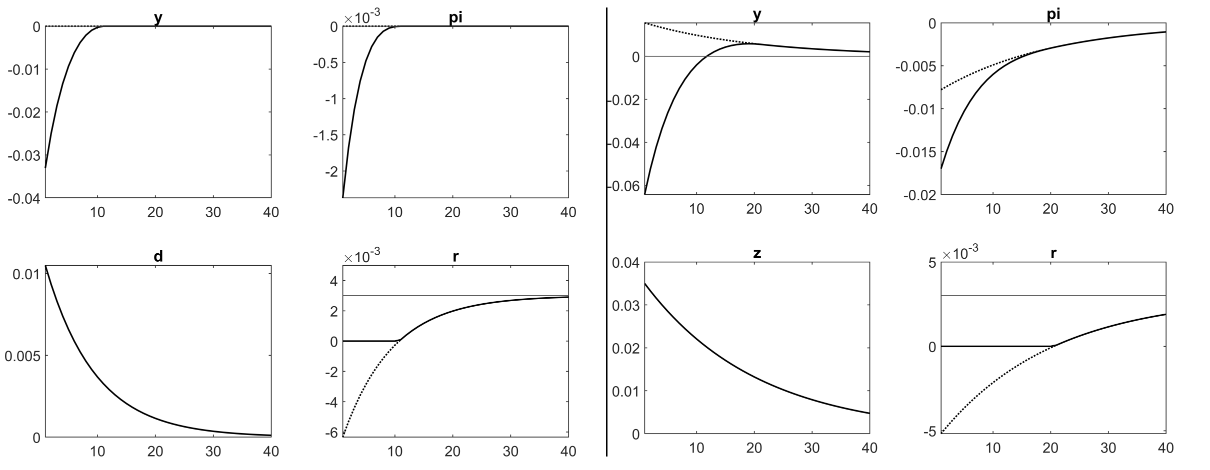


Figure 1: Sample output from running "dynareOBC bbw2016.mod ShockScale=5".

The left 4 panels show the response to a 5 standard deviation demand shock. The right 4 panels show the response to a 5 standard deviation productivity shock. All variables are in logarithms.

In all cases, the dotted lines show the path the economy would have followed were it not for the ZLB.

DynareOBC always also outputs diagnostic information about the model. For example, for this model, DynareOBC outputs the following, after it has made its final internal call to Dynare. Here we have made the most important lines bold to highlight them, and we have removed some additional white space:

Beginning to solve the model.

Solving the model for specific parameters.

Saving NLMA parameters.

Retrieving IRFs to shadow shocks.

Preparing normalized sub-matrices.

Largest P-matrix found with a simple criterion included elements up to horizon 32 periods.

The search for solutions will start from this point.

Pre-calculating the augmented state transition matrices and possibly conditional covariances.

Performing initial checks on the model.

M is an S matrix, so the LCP is always feasible. This is a necessary condition for there to always be a solution.

varsigma bounds (positive means M is an S matrix):

6.7600451299659 6.76004512996604

sum of y from the alternative problem (zero means M is an S matrix):

0

Skipping tests of feasibility with infinite T (TimeToEscapeBounds).

To run them, set FeasibilityTestGridSize=INTEGER where INTEGER>0.

Skipping further P tests, since we have already established that M is a P-matrix.

The M matrix with T (TimeToEscapeBounds) equal to 32 is a P-matrix. There is a unique solution to the model, conditional on the bound binding for at most 32 periods.

This is a necessary condition for M to be a P-matrix with arbitrarily large T (TimeToEscapeBounds).

A weak necessary condition for M to be a P-matrix with arbitrarily large T (TimeToEscapeBounds) is satisfied.

Discovering and testing the installed MILP solver.

Found working solver: GUROBI

Forming optimizer.

Preparing to simulate the model.

Simulating IRFs.

Cleaning up.

We see that DynareOBC's fast default diagnostics already identified that the M matrix for this model was a P-matrix and an S-matrix, as well as providing some weak evidence that M is a P-matrix for arbitrarily high T.

Note that DynareOBC refers to T as "TimeToEscapeBounds". This is the name DynareOBC gives to the command line option to control the size of the linear complementarity problems DynareOBC solves internally. To see why this may be necessary, try running the command:

dynareOBC bbw2016.mod ShockScale=10

Now DynareOBC does not complete successfully. Instead it reports:

Error using SolveBoundsProblem (line 241)

Impossible problem encountered. Try increasing TimeToEscapeBounds, or reducing the magnitude of shocks.

To avoid this problem, we just need to follow the advice of the error message and run with a higher value for "TimeToEscapeBounds". For example, if we run:

dynareOBC bbw2016.mod ShockScale=10 TimeToEscapeBounds=64

then DynareOBC completes successfully. In this case the response to the productivity shock stays at the ZLB for more than 32 periods, which is significant as 32 is the default number of periods for "TimeToEscapeBounds".

"TimeToEscapeBounds" and "ShockScale" are two of DynareOBC's command line options. There is a full list of these options in the "ReadMe.pdf" contained in DynareOBC's root directory, along with details on what each option does. Since the full list of options may be somewhat bewildering though, we conclude this getting started guide with details of those options most relevant to the analysis of a model's properties and those which impact perfect foresight simulation. Note that all options accepting a number must be entered without a space between the name of the option, the equals sign and the number.

- TimeToEscapeBounds=INTEGER (default: 32)

  The number of periods after which the model is expected to be away from any occasionally binding constraints. If there is no solution which finally escapes within this time, DynareOBC will produce an error.

- TimeToReturnToSteadyState=INTEGER (default: 64)

  The number of periods in which to verify that the constraints are not being violated.

- ReverseSearch

  By default, DynareOBC finds a solution in which the last period at the bound is as soon as possible. This option makes DynareOBC find a solution in which the last period at the bound is as remote as possible, subject to being less than the longest horizon (i.e. TimeToEscapeBounds).

- FullHorizon

  By default, DynareOBC finds a solution in which the last period at the bound is as soon as possible. This option makes DynareOBC just solve the bounds problem at the longest horizon.

- Omega=FLOAT (default: 1000)

  The tightness of the constraint on the news shocks. If this is large, solutions with news shocks close to zero will be returned when there are multiple solutions. It is often helpful to combine this option with FullHorizon so that DynareOBC does not just choose the solution which escapes the bound first.

- SkipFirstSolutions=INTEGER (default: 0)

If this is greater than 0, then DynareOBC ignores the first INTEGER solutions it finds, unless no other solutions are found, in which case it takes the last found one. Thus, without ReverseSearch, this tends to find solutions at the bound for longer. With ReverseSearch, this tends to find solutions at the bound for less time.

- FeasibilityTestGridSize=INTEGER (default: 0)

  Specifies the number of points in each of the two axes of the grid on which a test of a sufficient condition for feasibility with $T = \infty$ is performed. Setting a larger number increases the chance of finding feasibility, but may be slow.

  If FeasibilityTestGridSize=0 then the test is disabled.

- SkipQuickPCheck

  Disables the "quick" check to see if the M matrix has any contiguous principal sub-matrices with non-positive determinants.

- PTest=INTEGER (default: 0)

  Runs a fast as possible test to see if the top INTEGERxINTEGER submatrix of M is a P-matrix. Set this to 0 to disable these tests.

- AltPTest=INTEGER (default: 0)

  Uses a slower, more verbose procedure to test if the top INTEGERxINTEGER submatrix of M is a P-matrix. Set this to 0 to disable these tests.

- FullTest=INTEGER (default: 0)

  Runs very slow tests to see if the top INTEGERxINTEGER submatrix of M is a $P_{(0)}$ and/or (strictly) semi-monotone matrix.

- UseVPA

  Enables more accurate evaluation of determinants using the symbolic toolbox.

- ShockScale=FLOAT (default: 1)

  Scale of shocks for IRFs. This allows the calculation of IRFs to shocks larger or smaller than one standard deviation.

- IRFsAroundZero

  By default at first order, IRFs are centred around the steady state. This option instead centres IRFs around 0.

Appendix B: Additional matrix properties and their relationships

The following definitions help us state our additional results:[1]

Definition 7 (Principal sub-matrix, Principal minor) For a matrix $M \in \mathbb{R}^{T\times T}$, the **principal sub-matrices** of M are the matrices $\left[M_{i,j}\right]_{i,j=k_1,\dots,k_S}$, where $S, k_1, \dots, k_S \in \{1, \dots, T\}, k_1 < k_2 < \cdots < k_S$, i.e. the **principal sub-matrices** of M are formed by deleting the same rows and columns. The **principal minors** of M are the determinants of M's principal sub-matrices.

Definition 8 (P(₀)-matrix) A matrix $M \in \mathbb{R}^{T\times T}$ is called a **P-matrix ($\mathbf{P_0}$-matrix)** if the principal minors of M are all strictly (weakly) positive.

Definition 9 (General positive (semi-)definite) A matrix $M \in \mathbb{R}^{T\times T}$ is called **general positive (semi-)definite** if $M + M'$ is positive (semi-)definite (p.(s.)d.).

Definition 10 ((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.

Definition 11 (Sufficient matrices) $M \in \mathbb{R}^{T\times T}$ is called **column sufficient** if M is a P_0-matrix, and for each principal sub-matrix $W := \left[M_{i,j}\right]_{i,j=k_1,\dots,k_S}$ 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 $\left[W_{i,j}\right]_{\substack{i=1,\dots,S\\ j=l_1,\dots,l_R}}$ are not a basis for the column space of W.[2] M is called **row sufficient** if M' is column sufficient. M is called **sufficient** if it is column and row sufficient.

Definition 12 ((Strictly) Copositive) A matrix $M \in \mathbb{R}^{T\times T}$ is called (**strictly**) **copositive** if $M + M'$ is (strictly) semi-monotone.[3]

Definition 13 (Adequate matrices) $M \in \mathbb{R}^{T\times T}$ is called **column adequate** if M is a P_0-matrix, and for each principal sub-matrix $W := \left[M_{i,j}\right]_{i,j=k_1,\dots,k_S}$ of M with zero determinant, the

[1] 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. For the original definitions, and the proofs of equivalence between the ones below and the originals, see Cottle, Pang & Stone (2009a) and Xu (1993).

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

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

columns of $\left[M_{i,j}\right]_{\substack{i=1,\ldots,T\\ j=k_1,\ldots,k_S}}$ are linearly dependent. M is called **row adequate** if M' is column adequate. M is called **adequate** if it is column adequate and row adequate.

Definition 14 (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 $My \gg 0$ ($My \geq 0$).[4]

Definition 15 ((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**).

For example, consider the $T = 3$ case with $M = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix}$. Then M is a P-matrix if and only if $M_{11} > 0$, $M_{22} > 0$, $M_{33} > 0$, $\det \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} > 0$, $\det \begin{bmatrix} M_{11} & M_{13} \\ M_{31} & M_{33} \end{bmatrix} > 0$, $\det \begin{bmatrix} M_{22} & M_{23} \\ M_{32} & M_{33} \end{bmatrix} > 0$ and $\det M > 0$.

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

Lemma 1 The following hold:

1) All general positive semi-definite matrices are copositive, sufficient and P$_0$.
2) All general positive definite matrices are P matrices.
3) P$_0$ includes skew-symmetric, general positive semi-definite, sufficient and P matrices.
4) All P$_0$-matrices, and all copositive matrices are semi-monotone, and all P-matrices, and all strictly copositive matrices are strictly semi-monotone.
5) All column (row) adequate matrices are column (row) sufficient.

A common intuition is that in models without state variables, M must be both a P matrix, and an S matrix. This is not true. There are even purely static models for which M is in neither of these classes, as we prove the following result in Appendix H.5. See also Corollary 2 from the main paper.

[4] These conditions may be rewritten as $\sup\{\varsigma \in \mathbb{R} | \exists y \geq 0 \text{ s.t. } \forall t \in \{1,\ldots,T\}, (My)_t \geq \varsigma \wedge y_t \leq 1\} > 0$, and $\sup\left\{\sum_{t=1}^{T} y_t \middle| y \geq 0, My \geq 0 \wedge \forall t \in \{1,\ldots,T\}, y_t \leq 1\right\} > 0$, respectively. As linear programming problems, these may be solved in time polynomial in T using the methods of e.g. Roos, Terlaky, and Vial (2006). (More precisely, they can be solved in time proportional to $T^{2.37}$ in the worst case, by the result of Jiang et al. (2020), using the current best bounds on the "ω" and "α" constants on which their results depend.) Alternatively, by Ville's Theorem of the Alternative (Cottle, Pang & Stone 2009b), M is not an S$_0$-matrix if and only if $-M'$ is an S-matrix.

Proposition 3 There is 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.

Appendix C: Supplemental results

This section starts by presenting additional uniqueness results. The next subsection gives further existence results. Then, we examine the properties of small LCPs (with $T = 1$ or $T = 2$) in more detail. The section finishes with a discussion on how to approach checking for existence and uniqueness in practice.

Appendix C.1: Uniqueness

The following corollary of Theorem 2 gives more easily verified necessary conditions for uniqueness.

Corollary 6 If for all $q \in \mathbb{R}^T$, the LCP (q, M) has a unique solution, then:

1. All of the principal sub-matrices of M are P-matrices, S-matrices and strictly semi-monotone. (Cottle, Pang & Stone 2009a)
2. M has a strictly positive diagonal. (Immediate from definition.)
3. All of the eigenvalues of M have complex arguments in the interval $\left(-\pi + \frac{\pi}{T}, \pi - \frac{\pi}{T}\right)$. (Fang 1989)

The following corollary of Theorem 2 gives more easily verified sufficient conditions for uniqueness.

Corollary 7 For an arbitrary matrix A, denote the spectral radius of A by $\rho(A)$, and its largest and smallest singular values by $\sigma_{\max}(A)$ and $\sigma_{\min}(A)$, respectively. Let $|A|$ be the matrix with $|A|_{ij} = |A_{ij}|$ for all i, j. Then, for any matrix $M \in \mathbb{R}^{T\times T}$, if there exist diagonal matrices $D_1, D_2 \in \mathbb{R}^{T\times T}$ with positive diagonals, such that $W := D_1 M D_2$ satisfies one of the following conditions, then for all $q \in \mathbb{R}^T$, the LCP (q, M) has a unique solution:

1. W is general positive definite. (Cottle, Pang & Stone 2009a)
2. W has a positive diagonal, and $\langle W\rangle^{-1}$ is a nonnegative matrix, where $\langle W\rangle$ is the matrix with $\langle W\rangle_{ij} = -|W_{ij}|$ for $i \neq j$ and $\langle W\rangle_{ii} = |W_{ii}|$. (Bai & Evans 1997)
3. $\rho(|I - W|) < 1$. (Li & Wu 2016)
4. $(I + W)'(I + W) - \sigma_{\max}(|I - W|)^2 I$ is positive definite. (Li & Wu 2016)

5. $\sigma_{\max}(|I - W|) < \sigma_{\min}(I + W)$. (Li & Wu 2016)
6. $\sigma_{\min}\left((I - W)^{-1}(I + W)\right) > 1$. (Li & Wu 2016)
7. $\sigma_{\max}\left((I + W)^{-1}(I - W)\right) < 1$. (Li & Wu 2016)
8. $\rho(\left|(I + W)^{-1}(I - W)\right|) < 1$. (Li & Wu 2016)

In our experience, whenever M is a P-matrix, it will usually satisfy one of these conditions when D_1 and D_2 are chosen so that all rows and columns of $|W|$ have maximum equal to 1, using the algorithm of Ruiz (2001).

We also have necessary conditions for uniqueness with arbitrary T. In particular:

Proposition 4 Given an otherwise linear model with an OBC, the limit, d_k of the k^{th} diagonal[5] of M with $T = \infty$ exists, is finite, and is computable in time polynomial in k and the number of state variables of the model. If for all finite T, M is a P-matrix, then for all $S > 0$, the $S \times S$ Toeplitz matrix with k^{th} diagonal d_k is a P_0-matrix.

The properties of the limits of the diagonals of M are established in Appendix H.4 as part of the proof of Proposition 2. The rest of the claim follows from the continuity of determinants.

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 5 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 & Stone 2009a)

Hence, by verifying that M is semi-monotone, we can reassure ourselves that 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. 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.

[5] We take diagonal indices to be increasing as one moves up and right in M.

Where there are multiple solutions, we might like to select one via some objective function. This is tractable when either the number of solutions is finite, or the solution set is convex:

Proposition 6 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 & Stone 2009a)

Proposition 7 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 & Stone 2009a)

Finally, conditions for uniqueness of the path of the bounded variable is given in the following proposition:

Proposition 8 There exists w such that for any solution y of the LCP (q, M), $q + My = w$ if and only if M is column adequate. (Cottle, Pang & Stone 2009a)

Appendix C.2: Existence

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

Proposition 9 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 & Stone 2009a; Väliaho 1986)

If either condition 1 or condition 2 of Proposition 9 is satisfied, then to check existence for any particular q, we only need to solve a linear programming problem. As this will be faster than solving the particular LCP, this may be helpful in practice. Moreover:

Proposition 10 The LCP (q, M) is solvable for all $q \in \mathbb{R}^T$, if at least one of the following conditions holds: (Cottle, Pang & Stone 2009a)

1. M is an S-matrix, and either condition 1 or 2 of Proposition 9 is satisfied.
2. M is copositive and non-degenerate.
3. M is a P-, a strictly copositive or strictly semi-monotone matrix.

If condition 1, 2 or 3 of Proposition 10 is satisfied, then 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. Finally, in the special case of nonnegative M matrices we can derive conditions for existence that are both necessary and sufficient:

Proposition 11 If M is a nonnegative matrix, then the LCP (q, M) is solvable for all $q \in \mathbb{R}^T$ if and only if M has a positive diagonal. (Cottle, Pang & Stone 2009a)

Appendix C.3: Small LCPs

LCPs of size 1

When $T = 1$, it is particularly easy to characterise the properties of LCPs. This amounts to considering the behaviour of an economy in which everyone believes there will be at most one period at the bound. In this case, y gives the "shock" to the bounded equation necessary to impose the bound, and M gives the contemporaneous response of the bounded variable to an unanticipated shock: i.e., in a ZLB context, M gives the initial jump in nominal interest rates following a standard monetary policy shock.

First, suppose that M (a scalar as $T = 1$ for now) is positive. Then, if $q > 0$, for any $y \geq 0$, $q + My > 0$, so by the complementary slackness condition, in fact $y = 0$. Conversely, if $q \leq 0$, then there is a unique y satisfying the complementary slackness condition given by $y = -\frac{q}{M} \geq 0$. Thus, with $M > 0$, there is always a unique solution to the $T = 1$ LCP. With $M = 0$, $q + My = q$, so a solution to the LCP exists if and only if $q \geq 0$. It will be unique providing $q > 0$ (by the complementary slackness condition), but when $q = 0$, any $y \geq 0$ gives a solution. Finally, suppose that $M < 0$. Then, if $q > 0$, there are precisely two solutions. The "standard" solution has $y = 0$, but there is an additional solution featuring a jump to the bound in which $y = -\frac{q}{M} > 0$. If $q = 0$, then there is a unique solution ($y = 0$) and if $q < 0$, then with $y \geq 0$, $q + My < 0$, so there is no solution at all. Hence, the $T = 1$ LCP already provides examples of cases of uniqueness, non-existence and multiplicity.

LCPs of size 2

We now consider the $T = 2$ special case, where we can again easily derive results from first principles. Recall that a solution $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ to the LCP $\left(\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \right)$ satisfies $y_1 \geq 0$, $y_2 \geq 0$, $q_1 + M_{11} y_1 + M_{12} y_2 \geq 0$, $q_2 + M_{21} y_1 + M_{22} y_2 \geq 0$, $y_1 (q_1 + M_{11} y_1 + M_{12} y_2) =$

0, and $y_2(q_2 + M_{21}y_1 + M_{22}y_2) = 0$. With two quadratics, there are up to four generic solutions, given by:

1) $y_1 = y_2 = 0$. Exists if $q_1 \geq 0$ and $q_2 \geq 0$.
2) $y_1 = -\frac{q_1}{M_{11}}, y_2 = 0$. Exists if $\frac{q_1}{M_{11}} \leq 0$ and $M_{11}q_2 \geq M_{21}q_1$.
3) $y_1 = 0, y_2 = -\frac{q_2}{M_{22}}$. Exists if $\frac{q_2}{M_{22}} \leq 0$ and $M_{22}q_1 \geq M_{12}q_2$.
4) $y_1 = \frac{M_{12}q_2 - M_{22}q_1}{M_{11}M_{22} - M_{12}M_{21}}, y_2 = \frac{M_{21}q_1 - M_{11}q_2}{M_{11}M_{22} - M_{12}M_{21}}$. Exists if $y_1 \geq 0$ and $y_2 \geq 0$.

Additionally, there are extra solutions in knife-edge cases:

5) If $q_1 = 0, M_{11} = 0$ and $q_2 \geq 0$ then any $y_1 \geq 0$ is a solution with $y_2 = 0$.
6) If $q_2 = 0, M_{22} = 0$ and $q_1 \geq 0$ then any $y_2 \geq 0$ is a solution with $y_1 = 0$.
7) If $q_1 = 0$, $q_2 = 0$, $M_{11}M_{22} = M_{12}M_{21}$, then any $y_1 \geq 0$ and $y_2 \geq 0$ with $M_{21}y_1 = -M_{22}y_2$ is a solution.
8) If $q_1 = 0, q_2 = 0, M_{11} = M_{12} = M_{21} = M_{22}$, then any $y_1 \geq 0$ and $y_2 \geq 0$ are a solution.

Appendix C.4: Checking the existence and uniqueness conditions in practice

The paper has presented many results, but the practical details of what one should test and in what order may still be unclear. Luckily, a lot of the decisions are automated by the author's DynareOBC toolkit, but we present a suggested testing procedure here in any case. This also serves to give an overview of our results and their limitations.

For checking feasibility and existence, the most powerful result is Proposition 2 and Corollary 4. If the lower bound from Proposition 2 is positive, for all sufficiently high T, the LCP is always feasible. If further conditions are satisfied for a given T, (see Proposition 9 and Proposition 10) then this guarantees existence for that particular T. However, since the additional conditions are sufficient and not necessary, in practice it may not be worth checking them, as we have never encountered a problem without a solution that was nonetheless feasible.

The construction of the bounds in Proposition 2 (and its proof in Appendix H.4) requires a choice of a $T > 0$. Finding a T for which Proposition 2 produces a positive lower bound on ς requires a bit of trial and error. T will need to be big enough that the asymptotic approximation is accurate, which usually requires T to be bigger than the time it takes for the model's dynamics to die out. However, if T is too large, then DynareOBC's conservative approach to handling numerical error means that it can be difficult to reject $\underline{\varsigma} = 0$. Usually though, an intermediary value for T can be found at which we can establish $\underline{\varsigma} > 0$, even with a conservative approach to numerical error.

For checking non-existence, Proposition 2 and Corollary 4 can still be useful, though in this case, it does not provide definitive proof of non-feasibility, due to inescapable numerical inaccuracies. For a particular T, we may test if M is not an S-matrix in time polynomial in T by solving a simple linear programming problem. If M is not an S-matrix, then by Proposition 1 and Corollary 3, there are some q for which there is no path which does not violate the bound in the first T periods. With T larger than the time it takes for the model's dynamics to die out, this provides further evidence of non-existence for arbitrarily large T. In any case, given that only having a solution that stays at the bound for 250 years is arguably as bad as having no solution at all, for medium scale models, we suggest to just check if M is an S-matrix with $T = 1000$.

For checking uniqueness vs multiplicity, it is important to remember that while we can prove uniqueness for a given finite T by proving that the M matrix is a P-matrix, once we have found one T for which M is not a P-matrix (so there are multiple solutions, by Theorem 2 and Corollary 1), we know the same is true for all higher T. If we wish to prove that there is a unique solution up to some horizon T, then the best approach is to begin by testing the sufficient conditions from Corollary 7, with our suggested D_1 and D_2. If none of these conditions pass, then it is probable that M is not a P-matrix. In any case, checking that an M which fails the conditions of Corollary 7 is a P-matrix for very large T may not be computationally feasible, though finding a counter-example usually is. However, for purely forward-looking or purely backward-looking models, Corollary 2 tells us all we need to know for any T.

If we wish to establish multiplicity (for models that are not purely forward-looking or purely backward-looking), then Corollary 6 provides a guide. It is trivial to check if M has any nonpositive elements on its diagonal, in which case it cannot be a P-matrix. We can also check whether the expression derived in Appendix H.4 for the limit of the diagonal of M is non-negative, which is a necessary condition for M to be a P-matrix for all large T (this is a special case of Proposition 4). It is also trivial to check the eigenvalue condition given in Corollary 6, and that M is an S-matrix. If none of these checks established that M is not a P-matrix, then a search for a principal sub-matrix with negative determinant is the obvious next step. It is sensible to begin by checking the contiguous principal sub-matrices.[6] These correspond to a single spell at the ZLB which is natural given that impulse responses in DSGE models tend to be single peaked. This is so reliable a diagnostic (and so fast) that DynareOBC reports it automatically for all models. Continuing, one could then check all the 2×2 principal sub-matrices, then the 3×3 ones, and so on. With T around the half-life of the model's dynamics, usually one of these tests will quickly produce the required counter-example. A similar search strategy can be used to rule out semi-monotonicity, implying multiplicity when away from the bound, by Proposition 5.

Given the computational challenge of verifying whether M is a P-matrix, without Corollary 7, it may be tempting to wonder if our results really enable one to accomplish anything that could not have been accomplished by a naïve brute force approach. For example, it has been suggested that given T and an initial state, one could check for multiple equilibria by considering all of the 2^T possible combinations of periods at which the model could be at the bound and testing if each guess is consistent with the model, following, for example, the solution algorithms of Fair and Taylor (1983) or Guerrieri & Iacoviello (2015). Since there are 2^T principal sub-matrices of M, it might seem likely that this will be computationally very similar to checking if M is a P-matrix. However, our uniqueness results are not conditional on q or the initial state, rather they give conditions under which there is a unique solution for any possible path that the economy would take in the absence of the bound. Thus, while the brute force approach may eventually tell you about uniqueness given an initial state, using our results, in a comparable amount of time you will learn whether there are multiple solutions for any possible q. A brute force approach to checking for all possible initial conditions would require one to solve a linear programming problem for each pair of possible sets of periods at the bound, of which there are $2^{2T-1} - 2^{T-1}$.[7] This is far more computationally demanding than our

[6] Some care must be taken though as checking the signs of determinants of large matrices is numerically unreliable.

[7] Given the periods in the constrained regime, the economy's path is linear in the initial state. Excepting knife edge cases of rank deficiency, any multiplicity must involve two paths each at the bound in a different set of periods. Consequently, a brute

approach, and becomes intractable for even very small T. Additionally, our approach is numerically more robust, allows the easy management of the effects of numerical error to avoid false positives and false negatives, and requires less work in each step. Finally, we stress that in most cases, thanks to Corollary 6 and Corollary 7, no such search of the sub-matrices of M is required under our approach, and a proof or counter-example may be produced in time polynomial in T, just as it may be when checking for existence with our results.

Appendix D: Formal treatment of our equivalence result

Appendix D.1: Problem set-ups

In the absence of occasionally binding constraints, calculating an impulse response or performing a perfect foresight simulation exercise in a linear DSGE model is equivalent to solving the following problem:

Problem 1 (Linear) Suppose that $x_0 \in \mathbb{R}^n$ is given. Find $x_t \in \mathbb{R}^n$ for $t \in \mathbb{N}^+$ such that:

1) $x_t \to \mu$ as $t \to \infty$,
2) for all $t \in \mathbb{N}^+$:

$$0 = A(x_{t-1} - \mu) + B(x_t - \mu) + C(x_{t+1} - \mu), \tag{5}$$

The absence of shocks here is without loss of generality. For suppose $0 = \hat{A}(\hat{x}_{t-1} - \mu) + \hat{B}(\hat{x}_t - \mu) + \hat{C}(\hat{x}_{t+1} - \mu) + \widehat{D}\varepsilon_t$, with $\hat{x}_t \to \hat{\mu}$ as $t \to \infty$, and that $\varepsilon_t = 0$ for $t > S$. Then, if we define:

$$x_t := \begin{bmatrix} \hat{x}_t \\ \varepsilon_{t+1} \\ \vdots \\ \varepsilon_{t+S} \end{bmatrix}, \mu := \begin{bmatrix} \hat{\mu} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$A := \begin{bmatrix} \hat{A} & \widehat{D} & 0 & \cdots & 0 \\ 0 & 0 & I & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & I \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, B := \begin{bmatrix} \hat{B} & 0 & \cdots & 0 \\ 0 & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I \end{bmatrix}, C := \begin{bmatrix} \hat{C} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

then we are left with a problem in the form of Problem 1 (Linear), with the extended initial condition:

force approach to finding multiplicity unconditional on the initial state is to guess two different sets of periods at which the economy is at the bound, then solve a linear programming problem to find out if there is a value of the initial state for which the regimes on each path agree with their respective guesses.

$$x_0 = \begin{bmatrix} \hat{x}_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_S \end{bmatrix},$$

and the extended terminal condition $x_t \to \mu$ as $t \to \infty$.

We make the following assumption throughout the paper and these appendices:

Assumption 1 For any given $x_0 \in \mathbb{R}^n$, Problem 1 (Linear) has a unique solution, which (without loss of generality) takes the form $x_t = (I - F)\mu + Fx_{t-1}$, for $t \in \mathbb{N}^+$, where $0 = A + BF + CFF$ (so $F = -(B + CF)^{-1}A$), and where the eigenvalues of F are strictly inside the unit circle.

Conditions (A') and (B) from Sims's (2002) generalisation of the standard Blanchard-Kahn (1980) conditions are necessary and sufficient for Assumption 1 to hold. Further, to avoid dealing specially with the knife-edge case of exact unit eigenvalues in the part of the model that is solved forward, here 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 non-linear model, and which is also necessary for the linear model to have a unique steady state:

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

We are interested in models featuring occasionally binding constraints. We will concentrate on models featuring a single ZLB type constraint in their first equation, which does not bind in steady state, and 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 Appendix D.3. 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 the identity, I, and so on. Then, from adding $x_{1,t}$ to both sides of the first equation within the system (5), then incorporating a max, we produce the system of interest:

Problem 2 (OBC) 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:

1) $x_t \to \mu$ as $t \to \infty$,
2) for all $t \in \mathbb{N}^+$:

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

3) $x_{1,t} > 0$ for $t > T$.

given:

Assumption 3 $\mu_1 > 0$, where μ_1 is the first element of μ.

Were it not for the max, this problem would be identical to Problem 1 (Linear), providing that Assumption 3 holds, as the existence of a $T \in \mathbb{N}$ such that $x_{1,t} > 0$ for $t > T$ is guaranteed by the fact that $x_{1,t} \to \mu_1$ as $t \to \infty$.

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

Problem 3 (News) 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:

1) $x_t \to \mu, y_t \to 0$, as $t \to \infty$,
2) for all $t \in \mathbb{N}^+$:

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

This is a version of Problem 1 (Linear) with a forcing process ("news") up to horizon T added to the first equation. We use this representation in which the forcing process enters via an augmented state to make clear that this is also a special case of Problem 1 (Linear). By construction, the value of $y_{i,t}$ gives the shock that in period t is expected to arrive in i periods. (To be clear: the first index of $y_{i,t}$ indexes over the elements of the vector $y_t \in \mathbb{R}^T$; the second index of $y_{i,t}$ indexes over periods.) Hence, as there is no uncertainty, $y_{t,0}$ gives the shock that will hit in period t, i.e. $y_{1,t-1} = y_{2,t-2} = \cdots = y_{t,0}$ for $t \le T$, and $y_{1,t-1} = 0$ for $t > T$. Thus, the first equation of the first block could be rewritten:

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

which is in the form of equation (4) from the main paper.

Appendix D.2: Relationships between the problems

Since $y_{1,t-1} = 0$ for $t > T$, by Assumption 1, $(x_{T+1} - \mu) = F(x_T - \mu)$. Now define $s_{T+1} := 0$. Then with $t = T$, we have that $(x_{t+1} - \mu) = s_{t+1} + F(x_t - \mu)$. 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:

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

i.e., if we define: $s_t := -(B + CF)^{-1}(Cs_{t+1} + I_{\cdot,1}y_{t,0})$, then $(x_t - \mu) = s_t + F(x_{t-1} - \mu)$. By induction then, this holds for all $t \in \{1, \dots, T\}$, establishing:[8]

Lemma 2 There is a unique solution to Problem 3 (News) that is linear in x_0 and y_0.

For future reference, let $x_t^{(3,k)}$ be the solution to Problem 3 (News) 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 (News) when $x_0 = \mu$ is given by:

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

Now, let $M \in \mathbb{R}^{T\times T}$ satisfy:

$$M_{t,k} = x_{1,t}^{(3,k)} - \mu_1, \qquad \forall t, k \in \{1, .., T\}, \tag{6}$$

i.e. M horizontally stacks the (column-vector) relative impulse responses of the first variable to the news shocks, with the first column giving the response to a contemporaneous shock, the second column giving the response to a shock anticipated by one period, and so on. Then, this result implies that for arbitrary x_0 and y_0, the path of the first variable in the solution to Problem 3 (News) is given by:

$$(x_{1,1:T})' = q + My_0, \tag{7}$$

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

where $q := \left(x_{1,1:T}^{(1)}\right)'$ and where $x_t^{(1)}$ is the unique solution to Problem 1 (Linear), for the given x_0, i.e. q is the path of the first variable in the absence of news shocks or bounds.[9] Since M is not a function of either x_0 or y_0, equation (7) gives a highly convenient representation of the solution to Problem 3 (News).

Now let $x_t^{(2)}$ be a solution to Problem 2 (OBC) given some 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 $T' \le T$. We seek to relate the solution to Problem 2 (OBC) with the one to Problem 3 (News) for an appropriate choice of y_0. First, for all $t \in \mathbb{N}^+$, let:

$$\hat{e}_t := -\left[I_{1,\cdot}\mu + A_{1,\cdot}\left(x_{t-1}^{(2)} - \mu\right) + \left(B_{1,\cdot} + I_{1,\cdot}\right)\left(x_t^{(2)} - \mu\right) + C_{1,\cdot}\left(x_{t+1}^{(2)} - \mu\right)\right],$$
$$e_t := \begin{cases} \hat{e}_t & \text{if } x_{1,t}^{(2)} = 0 \\ 0 & \text{if } x_{1,t}^{(2)} > 0 \end{cases}, \tag{8}$$

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 that by the definition of Problem 2 (OBC), $e_t \ge 0$ and $x_{1,t}^{(2)} e_t = 0$, for all $t \in \mathbb{N}^+$.

Now, from the definition of Problem 2 (OBC), 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, if we define $\tilde{s}_{T+1} := 0$, then with $t = T$, $\left(x_{t+1}^{(2)} - \mu\right) = \tilde{s}_{t+1} + F\left(x_t^{(2)} - \mu\right)$. 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:

$$\left(x_t^{(2)} - \mu\right) = F\left(x_{t-1}^{(2)} - \mu\right) - (B + CF)^{-1}\left(C\tilde{s}_{t+1} + I_{\cdot,1}e_t\right),$$

i.e., if we define: $\tilde{s}_t := -(B + CF)^{-1}\left(C\tilde{s}_{t+1} + I_{\cdot,1}e_t\right)$, 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 (News) is started with $x_0 = x_0^{(2)}$ and $y_0 = e_{1:T}'$, then $x_t^{(2)}$ solves Problem 3 (News). Conversely, if $x_t^{(2)}$ solves Problem 3 (News) for some y_0, then from the laws of motion of x_t

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

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 establishes the following result:

Lemma 3 For any solution, $\left(T, x_t^{(2)}\right)$ to Problem 2 (OBC):

1) With $e_{1:T}$ as defined in equation (8), $e_{1:T} \geq 0$, $x_{1,1:T}^{(2)} \geq 0$ 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 (News) with $x_0 = x_0^{(2)}$ and $y_0 = e'_{1:T}$.
3) If $x_t^{(2)}$ solves Problem 3 (News) with $x_0 = x_0^{(2)}$ and with some y_0, then $y_0 = e'_{1:T}$.

To use the easy solution to Problem 3 (News) to assist us in solving Problem 2 (OBC) just requires one more result. In particular, we need to show that if $y_0 \in \mathbb{R}^T$ 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 (News) when started at x_0, y_0, then $x_t^{(3)}$ must also be a solution to Problem 2 (OBC).

So, 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 (News) 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 (OBC). I.e., we must prove that for all $t \in \mathbb{N}^+$:

$$x_{1,t}^{(3)} = \max\left\{0, I_{1,\cdot}\mu + A_{1,\cdot}\left(x_{t-1}^{(3)} - \mu\right) + \left(B_{1,\cdot} + I_{1,\cdot}\right)\left(x_t^{(3)} - \mu\right) + C_{1,\cdot}\left(x_{t+1}^{(3)} - \mu\right)\right\}, \quad (9)$$
$$0 = A_{-1,\cdot}\left(x_{t-1}^{(3)} - \mu\right) + B_{-1,\cdot}\left(x_t^{(3)} - \mu\right) + C_{-1,\cdot}\left(x_{t+1}^{(3)} - \mu\right).$$

By the definition of Problem 3 (News), the latter equation must hold with equality. Hence, we just need to prove that equation (9) 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 (News):

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

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 (News). Thus:

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

Consequently, equation (9) holds in this case too, completing the proof.

Together with Lemma 2, Lemma 3, and our representation of the solution of Problem 3 (News) from equation (7), this completes the proof of the following key theorem:

Theorem 1 (Restated) The following hold:

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

Appendix D.3: Generalizations

It is straightforward to generalise the results of this paper to less restrictive otherwise linear models with occasionally binding constraints.

Firstly, if the constraint is on a variable other than $x_{1,t}$, or in another equation than the first, then all of the results go through as before, just by relabelling and rearranging. Furthermore, if the constraint takes the form of $z_{1,t} = \max\{z_{2,t}, z_{3,t}\}$, where $z_{1,t}$, $z_{2,t}$ and $z_{3,t}$ are linear expressions in the contemporaneous values, lags and leads of x_t, then, assuming without loss of generality that $z_{3,\cdot} > z_{2,\cdot}$ in steady state, we have that $z_{1,t} - z_{2,t} = \max\{0, z_{3,t} - z_{2,t}\}$. Hence, adding a new auxiliary variable $x_{n+1,t}$, with the associated equation $x_{n+1,t} = z_{1,t} - z_{2,t}$, and replacing the constrained equation with $x_{n+1,t} = \max\{0, z_{3,t} - z_{2,t}\}$, we have a new equation in the form covered by our results. Moreover, if rather than a max we have a min, we just use the fact that if $z_{1,t} = \min\{z_{2,t}, z_{3,t}\}$, then $-z_{1,t} = \max\{-z_{2,t}, -z_{3,t}\}$, which is covered by the generalisation just established. The easiest encoding of the complementary slackness conditions, $z_t \geq 0$, $\lambda_t \geq 0$ and $z_t\lambda_t = 0$, is $0 = \min\{z_t, \lambda_t\}$, which is of this form.

To deal with multiple occasionally binding constraints, we use the representation from Holden and Paetz (2012). Suppose there are c constrained variables in the model. For $a \in \{1, \dots, c\}$, let $q^{(a)}$ be the path of the a^{th} constrained variable in the absence of all constraints. For

$a, b \in \{1, \ldots, c\}$, let $M^{(a,b)}$ be the matrix whose k^{th} column is the impulse response of the a^{th} constrained variable to magnitude 1 news shocks at horizon $k-1$ to the equation defining the b^{th} constrained variable. For example, if $c = 1$ so there is a single constraint, then we would have that $M^{(1,1)} = M$ as defined in equation (6). Finally, let:

$$q := \begin{bmatrix} q^{(1)} \\ \vdots \\ q^{(c)} \end{bmatrix}, \qquad M := \begin{bmatrix} M^{(1,1)} & \cdots & M^{(1,c)} \\ \vdots & \ddots & \vdots \\ M^{(c,1)} & \cdots & M^{(c,c)} \end{bmatrix},$$

and let y be a solution to the LCP (q, M). Then the vertically stacked paths of the constrained variables in a solution which satisfies these constraints is given by $q + My$, and Theorem 1 (Restated) goes through as before.

Appendix E: Example applications to New Keynesian models

In the first subsection here, we examine the simple Brendon, Paustian & Yates (BPY) (2013) model, before going on to consider a variant of it with price targeting, which we show to produce determinacy. In the BPY (2013) model, multiplicity and non-existence stem from a response to growth rates in the Taylor rule. However, we do not want to give the impression that multiplicity and non-existence are only caused by such a response, or that they are only a problem in carefully constructed theoretical examples. Thus, in Appendix E.2, 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. The next subsection shows that these conclusions also carry through to the posterior-modes of the Smets & Wouters (2003; 2007) models, and discusses the plausibility of self-fulfilling jumps to the ZLB. Finally, in Appendix E.4 we look at whether a simple model is determinate around the deflationary steady state.

Appendix E.1: Variants of the Brendon, Paustian & Yates (BPY) (2013) model

Brendon, Paustian & Yates (2013; 2019), is one of the most relevant pieces of prior work for the current paper. Like us, these authors examined perfect foresight equilibria of NK models with terminal conditions. In BPY (2013), the authors show analytically that in a very simple NK model, featuring a response to the growth rate in the Taylor rule, there are multiple perfect-foresight equilibria when all 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. Furthermore, the authors show numerically that in some select other models, there are multiple perfect-foresight equilibria when the economy begins at the steady state, and all agents believe that the economy will jump to the bound, remain there for some number of periods, before leaving it endogenously, after which they believe they will never hit the bound again. BPY (2019) extends these results to fully non-linear rational expectations equilibria in certain simple NK models.

Relative to these authors, we focus more on providing general theoretical results, which permit numerical analysis (for the otherwise linear, perfect foresight case) that is both more robust and less restrictive. This robustness and generality is crucial in showing multiplicity even in simple NK models, with entirely standard Taylor rules. For example, whereas in an intermediary working paper BPY (2016) write that price-dispersion "does not have a strong enough impact on equilibrium allocations for the sort of propagation that we need", we show that the presence of price dispersion is sufficient for multiplicity. Our general results are also crucial for allowing us to show uniqueness under a price level target.

The simple Brendon, Paustian & Yates (BPY) (2013) model

Brendon, Paustian & Yates (2013), provide a simple New Keynesian model that we can use to illustrate the possibility for multiplicity in NK models. Its equations follow:[10]

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

where $x_{i,t}$ is the nominal interest rate, $x_{y,t}$ is the output gap, $x_{\pi,t}$ is inflation, and $\beta \in (0,1)$, $\gamma, \sigma, \alpha_{\Delta y} \in (0, \infty)$, $\alpha_\pi \in (1, \infty)$ are parameters. The model's only departure from the textbook three equation NK model is the presence of an output growth rate term in the Taylor rule. This introduces an endogenous state variable in a tractable manner. In Appendix H.8, below, we prove the following:

[10] An implementation of this model is contained within DynareOBC in the file "Examples/BrendonPaustianYates2013/BPYModel.mod".

Proposition 12 The BPY model is in the form of Problem 2 (OBC), and satisfies Assumptions 1, 2 and 3. With $T = 1$, $M < 0$ ($M = 0$) if and only if $\alpha_{\Delta y} > \sigma\alpha_\pi$ ($\alpha_{\Delta y} = \sigma\alpha_\pi$).

Hence, by Theorem 1 (Restated) , when all agents believe the bound will be escaped after at most one period, if $\alpha_{\Delta y} < \sigma\alpha_\pi$, the model has a unique solution for all q, i.e. no matter what the nominal interest rate would be that period were no ZLB. If $\alpha_{\Delta y} = \sigma\alpha_\pi$, then the model has a unique solution whenever $q > 0$, infinitely many solutions when $q = 0$, and no solutions leaving the ZLB after one period when $q < 0$. Finally, if $\alpha_{\Delta y} > \sigma\alpha_\pi$ then the model has two solutions when $q > 0$, one solution when $q = 0$ and no solution escaping the ZLB next period when $q < 0$.

The mechanism here is as follows. The stronger the response to the growth rate, the more persistent is output, as the monetary rule implies additional stimulus if output was high last period. Suppose then that there was an unexpected positive shock to nominal interest rates. Then, due to the persistence, this would lower not just output and inflation today, but also output and inflation next period. With low expected inflation, real interest rates are high, giving consumers an additional reason to save, and thus further lowering output and inflation this period and next. With sufficiently high $\alpha_{\Delta y}$, this additional amplification is so strong that nominal interest rates fall this period, despite the positive shock, explaining why M may be negative.[11] Now, consider varying the magnitude of the original shock. For a sufficiently large shock, interest rates would hit zero. At this point, there is no observable evidence that a shock has arrived at all, since the ZLB implies that given the values of output and inflation, nominal interest rates should be zero even without a shock. Such a jump to the ZLB must then be a self-fulfilling prophecy. Agents expect low inflation, so they save, which, thanks to the monetary rule, implies low output tomorrow, rationalising the expectations of low inflation.

[11] Note that this cannot happen in the canonical 3 equation NK model in which the central bank responds to the output gap, not output growth. For, without state variables, in the period after the shock's arrival, inflation will be at steady state. Thus, in the period of the shock, real interest rates move one for one with nominal interest rates. Were the positive shock to the nominal interest rate to produce a fall in its level, then the Euler equation would imply high consumption today, also implying high inflation today via the Phillips curve. But, with consumption, inflation, and the shock all positive, the nominal interest rate must be above steady state, contradicting our assumption that it had fallen.

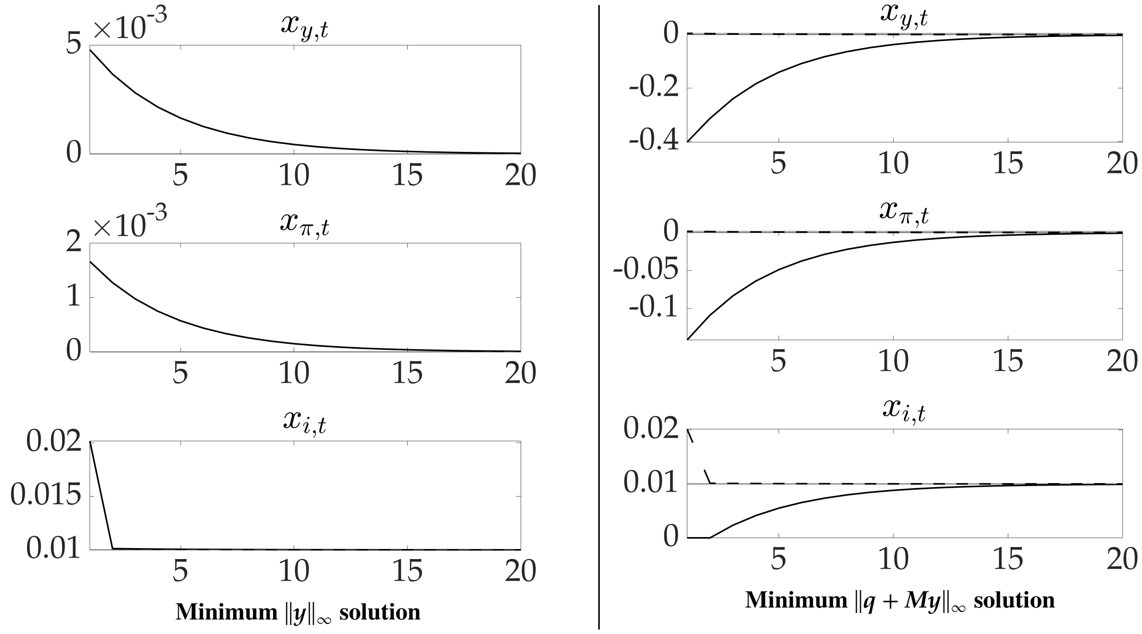


Figure 2: Alternative solutions following a magnitude 1 impulse to ε_t in the BPY model.

The dashed line in the right plot repeats the left plot, for comparison.

We finish this subsection with an example of multiplicity in the BPY (2013) model. This serves to illustrate the potential economic consequences of multiplicity in NK models. We present impulse responses to a shock to the Euler equation under two different solutions. With the shock added to the Euler equation, it now takes the form:

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

The other two equations of the BPY model equations remain as they were given above. We take the parameterisation $\sigma = 1$, $\beta = 0.99$, $\gamma = \frac{(1-0.85)(1-\beta(0.85))}{0.85}(2+\sigma)$, 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 dashed line there 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.[12]

[12] The plots in Figure 2 may be generated by navigating to the "Examples/BrendonPaustianYates2013" folder within DynareOBC, and then running "GeneratePlots".

The BPY model with shadow interest rate persistence

We showed that if $\alpha_{\Delta y} > \sigma\alpha_\pi$ in the BPY (2013) model, then with $T = 1, M < 0$. When $T > 1$, this implies that 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.

One obvious possibility is to consider models with either persistence in the interest rate, or persistence in the "shadow" rate that would hold were it not for the ZLB. Following BPY (2013), 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:[13]

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

It is easy to verify that this may be put in the form of Problem 2 (OBC), 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 $T = 4$. 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 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 ZLB episode over short-horizons.

Further evidence that the long-horizon behaviour is the same as in the model without persistence is provided by the fact that with $T = 20$, $\alpha_\pi = 1.5$ and $\alpha_{\Delta y} = 1.05$,[14] M is a P-

---

[13] An implementation of this model is contained within DynareOBC in the file: "Examples/BrendonPaustianYates2013/BPYModelPersistent.mod".

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

matrix. Moreover, from Proposition 2 with $T = 50$, we have that $\varsigma > 6.385 \times 10^{-8}$, so the model always has a feasible path, in the sense of Definition 6 (Feasibility), by Corollary 4.[15]

On the other hand, with $T = 200$, $\alpha_\pi = 1.5$ and $\alpha_{\Delta y} = 1.51$, then M is not an S-matrix,[16] meaning that for all sufficiently large T, M is not a P-matrix, so there are sometimes multiple solutions. Additionally, from Proposition 2 with $T = 200$, $\varsigma \leq 0 +$ numerical error, meaning that it is likely that the model does not have a solution for all possible paths of $x_{i,t}$.[17]

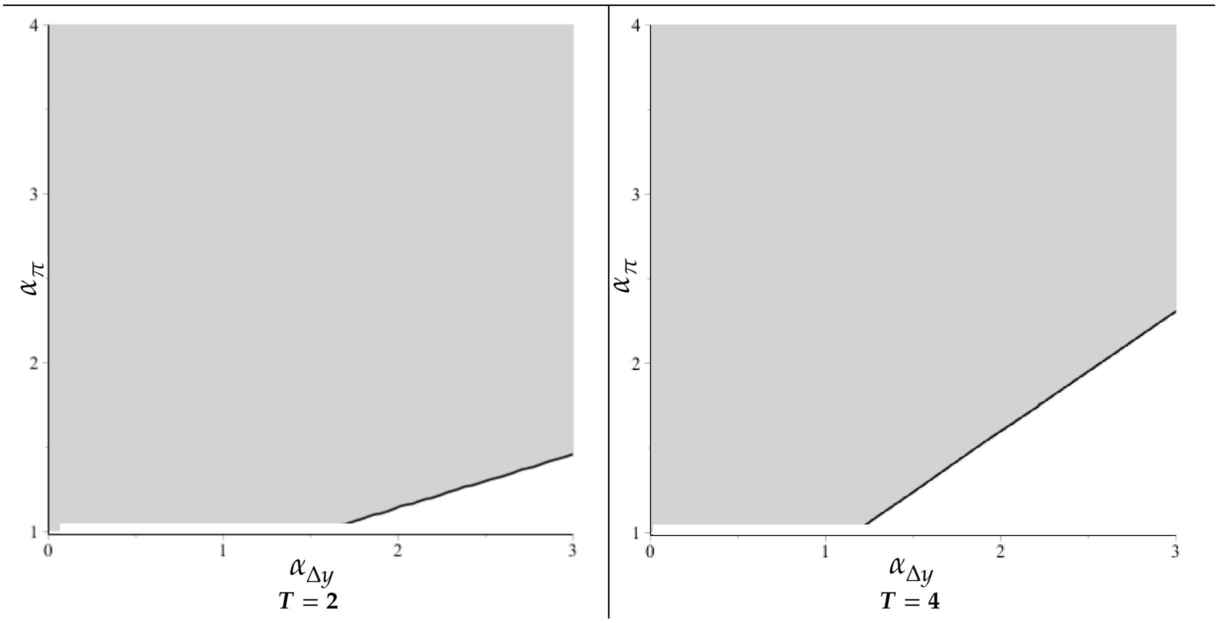


Figure 3: Regions in which M is a P-matrix (shaded grey) or a P_0-matrix (shaded grey, plus the black line), when $T = 2$ (left) or $T = 4$ (right).[18]

The BPY model with price level targeting

We may also introduce persistence in shadow interest rates by setting:

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

[15] This result is one of those produced by the "GenerateDeterminacyResults" script within the "Examples/BrendonPaustianYates2013" folder of DynareOBC.

[16] This was verified a second way by checking that $-M'$ was an S_0-matrix, as discussed in Footnote 4.

[17] These results are also among those produced by the "GenerateDeterminacyResults" script within the "Examples/BrendonPaustianYates2013" folder of DynareOBC.

[18] Code to generate this plot is contained within the Maple worksheet: "Examples/BrendonPaustianYates2013/AnalyticResults.mw".

where $x_{i,t} = \max\{0, x_{d,t}\}$. If the second bracketed term was multiplied by $(1-\rho)$, then this would be entirely standard, however as written here, 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 $\alpha_{\Delta y}, \alpha_{\pi} \in (0, \infty)$.

In particular, given the model:[19]

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

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

$$x_{p,t} - x_{p,t-1} = \beta \mathbb{E}_t x_{p,t+1} - \beta x_{p,t} + \gamma x_{y,t},$$

we prove in Appendix H.9, below, that the following proposition holds:

Proposition 13 The BPY model with price targeting is in the form of Problem 2 (OBC), and satisfies Assumptions 1, 2 and 3. 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 always has a feasible path, in the sense of Definition 6 (Feasibility). Given that $d_0 > 0$ for this model, and that for $T = 1000$, M is a P-matrix by our sufficient conditions from Corollary 7, this is strongly suggestive of the existence of a unique solution for any q and for arbitrarily large T.[20]

[19] An implementation of this model is contained within DynareOBC in the file "Examples/BrendonPaustianYates2013/BPYModelPriceLevelTargeting.mod".

[20] These results are also among those produced by the "GenerateDeterminacyResults" script within the "Examples/BrendonPaustianYates2013" folder of DynareOBC.

Appendix E.2: The linearized Fernández-Villaverde et al. (2015) model

The discussion of the BPY (2013) model 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 basic NK models without indexation to a positive steady-state inflation rate by non-optimising firms (and hence price dispersion in the steady state), still imply multiple equilibria in some states of the world (i.e. for some q) and no solutions in others, even with extremely aggressive monetary responses to inflation and without any monetary response to output growth. Price level targeting again fixes these problems though.

We show these results in the Fernández-Villaverde et al. (2015) 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).[21] The model's equilibrium conditions follow:

$$\frac{1}{C_t} = R_t \mathbb{E}_t \left[\frac{\beta_{t+1}}{\Pi_{t+1} C_{t+1}}\right] \qquad (*)$$

$$\psi L_t^{\vartheta} C_t = W_t$$

$$\varepsilon X_{1,t} = (\varepsilon - 1) X_{2,t}$$

$$X_{1,t} = \frac{Y_t}{C_t}\frac{W_t}{A_t} + \theta \mathbb{E}_t \beta_{t+1} \Pi_{t+1}^{\varepsilon} X_{1,t+1} \qquad (*)$$

$$X_{2,t} = \Pi_t^* \left(\frac{Y_t}{C_t} + \theta \mathbb{E}_t \beta_{t+1} \frac{\Pi_{t+1}^{\varepsilon-1}}{\Pi_{t+1}^*} X_{2,t+1}\right) \qquad (*)$$

$$\log R_t = \max\left\{0, \log R + \phi_\pi \log\left(\frac{\Pi_t}{\Pi}\right) + \phi_y \log\left(\frac{Y_t}{Y}\right) + \sigma_m \varepsilon_{m,t}\right\}$$

$$G_t = S_t Y_t$$

$$1 = \theta \Pi_t^{\varepsilon-1} + (1-\theta) {\Pi_t^*}^{1-\varepsilon}$$

$$\nu_t = \theta \Pi_t^{\varepsilon} \nu_{t-1} + (1-\theta) {\Pi_t^*}^{-\varepsilon} \qquad (*)$$

$$C_t + G_t = Y_t = \frac{A_t}{\nu_t} L_t$$

$$\log \beta_t = (1-\rho_\beta)\log\beta + \rho_\beta \log \beta_{t-1} + \sigma_\beta \varepsilon_{\beta,t}$$

$$\log A_t = (1-\rho_A)\log A + \rho_A \log A_{t-1} + \sigma_A \varepsilon_{A,t}$$

---

[21] An implementation of this model is contained within DynareOBC in the file "Examples/FernandezVillaverdeEtAl2015/NK.mod".

$$\log S_t = (1-\rho_S)\log S + \rho_S \log S_{t-1} + \sigma_S \varepsilon_{S,t}$$

Welfare in the model in period t is given by:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=0}^{s} \beta_{t+s}\right]\left[\log C_{t+s} - \frac{\psi}{1+\vartheta} L_t^{1+\vartheta}\right].$$

After substitutions, the model can be reduced to just the four non-linear equations marked with $(*)$ above (plus the three shock laws of motion) which are functions of gross inflation, Π_t, labour supply, L_t, price dispersion, ν_t, and an auxiliary variable introduced from the firms' price-setting first order condition, $X_{1,t}$, (plus the shocks). Of these variables, only price dispersion enters with a lag. We linearize the model around its steady state, and then reintroduce the "max" operator which linearization removed from the Taylor rule.[22] All parameters are set to the values given in Fernández-Villaverde et al. (2015). There is no response to output growth in the Taylor rule, so any multiplicity cannot be a consequence of the mechanism highlighted by BPY (2013).

For this model, numerical calculations reveal that with $T \leq 14$, M is a P-matrix. However, with $T \geq 15$, M is not a P matrix, and thus there are certainly some states of the world (some q) in which the model has multiple solutions. Furthermore, with $T = 1000$, our upper bound on ς from Proposition 2 implies that $\varsigma \leq 0 +$ numerical error, suggesting that the model does not have a solution for all possible paths of interest rates.[23]

To make the mechanism behind these results clear, we will compare the Fernández-Villaverde et al. (2015) model to an altered version of it with full indexation to steady-state inflation of prices that are not set optimally. To a first order approximation, the model with full indexation never has any price dispersion, and thus has no endogenous state variables. It is thus a purely forwards looking model, and so it is perhaps unsurprising that it should have a unique equilibrium given a terminal condition, even in the presence of the ZLB.

[22] Before linearization, we transform the model's variables so that the transformed variables 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 $\theta^{\frac{1}{1-\varepsilon}}$. Thus, we work with a logit transformation of inflation over $\theta^{\frac{1}{1-\varepsilon}}$.

[23] These results are among those that may be generated by running "GenerateDeterminacyResults" within the "Examples/FernandezVillaverdeEtAl2015" directory of DynareOBC.

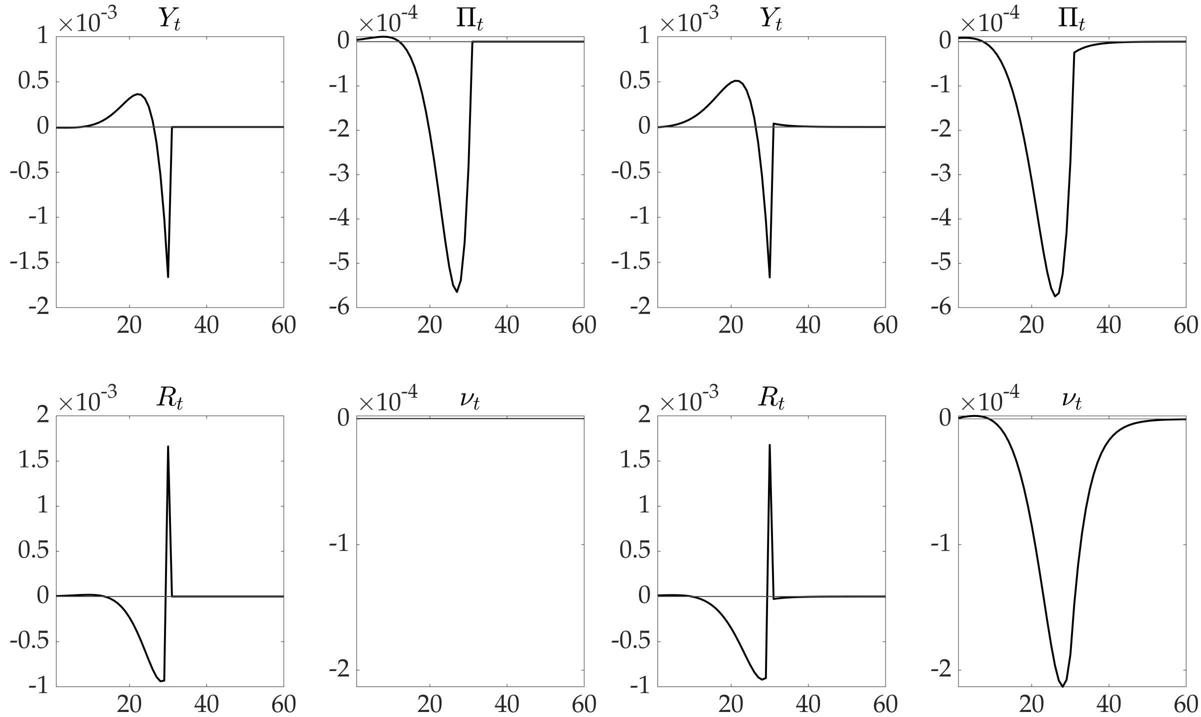


Figure 4: Impulse responses to a shock announced in period 1, but hitting in period 30, in basic New Keynesian models with (left 4 panels) and without (right 4 panels) indexation to steady-state inflation.

All variables are in logarithms. In both cases, the model and parameters are taken from Fernández-Villaverde et al. (2015), the only change being the addition of complete price indexation to steady-state inflation for non-updating firms in the left hand plots.

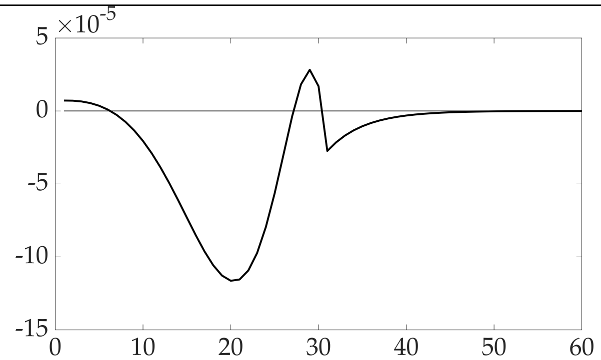


Figure 5: Difference between the IRFs of nominal interest rates from the two models shown in Figure 4.

Negative values imply that nominal interest rates are lower in the model without indexation.

In Figure 4 we plot the impulse responses of first order approximations to both models to a shock to nominal interest rates that is announced in period one but that does not hit until period thirty.[24] For both models, the shape is similar, however, in the model without indexation,

[24] This figure and the following ones in this subsection may be generated by running "GeneratePlots" within the "Examples/FernandezVillaverdeEtAl2015" directory of DynareOBC.

the presence of price dispersion reduces inflation both before and after the shock hits. This is because the predicted fall in inflation compresses the price distribution, reducing dispersion, and thus reducing the number of firms making large adjustments. The fall in price dispersion also increases output, due to lower efficiency losses from miss-pricing. However, the effect on interest rates is dominated by the negative inflation effect, as the Taylor-rule coefficient on output cannot be too high if there is to be determinacy.[25] For reference, the difference between the IRFs of nominal interest rates in each model is plotted in Figure 5, making clear that interest rates are on average lower following the shock in the model without indexation.

Remarkably, this small difference in the impulse responses between models is enough that the linearized model without indexation has multiple equilibria given a ZLB, but the linearized model with full indexation is determinate. This illustrates just how fragile is the uniqueness in the linearized purely forward-looking model. Informally, what is needed for multiplicity is that the impulse responses to positive news shocks to interest rates are sufficiently negative for a sufficiently high amount of time that a linear combination of them could be negative in every period in which a shock arrives. Here, price dispersion is providing the required additional reduction to nominal interest rates following a news shock.

We illustrate how multiplicity emerges in the model without indexation by showing, in Figure 6, the construction of an additional equilibrium which jumps to the ZLB for seventeen quarters.[26] If the economy is to be at the bound for seventeen quarters, then for those seventeen quarters, the nominal interest rate must be higher than it would be according to the Taylor rule, meaning that we need to consider seventeen endogenous news shocks, at horizons from zero to sixteen quarters into the future. The impulse responses to unit shocks of this kind are shown in the leftmost plot. Each impulse response has broadly the same shape as the one shown for nominal interest rates in the right of Figure 4. The central figure plots the same impulse

[25] One might think the situation would be different if the response to output was high enough that the rise in output after the shock produced a rise in interest rates. However, as observed by Ascari and Ropele (2009), the determinacy region is smaller in the presence of price dispersion than would be suggested by the Taylor criterion. Numerical experiments suggest that in all the determinate region, interest rates are below steady state following the shock.

[26] Seventeen quarters was the minimum span for which an equilibrium of this form could be found.

responses again, but now each line is scaled by a constant so that their sum gives the line shown in black in the rightmost plot. In this rightmost plot, the red line gives the ZLB's location, relative to steady state, thus the combined impulse response spends seventeen quarters at the ZLB before returning to steady state. Since there are only "news shocks" in the periods in which the economy is at the ZLB, this gives a perfect foresight rational expectations equilibrium which makes a self-fulfilling jump to the ZLB.

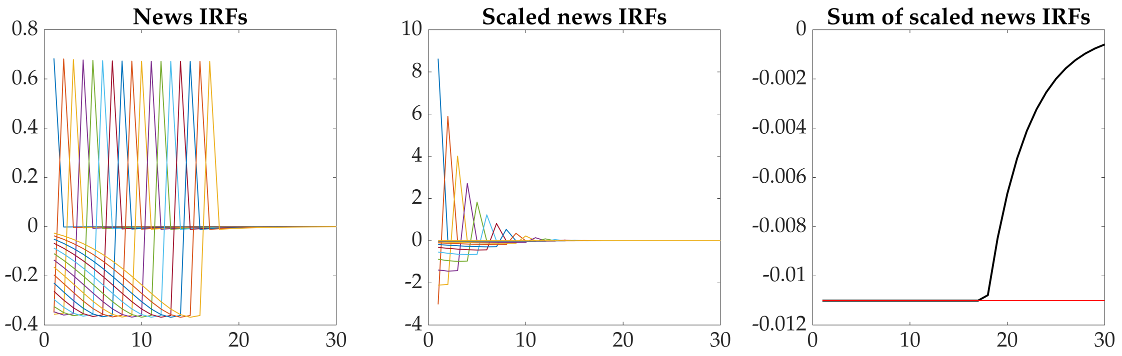


Figure 6: Construction of multiple equilibria in the Fernández-Villaverde et al. (2015) model.

The left plot shows the IRFs to news shocks arriving zero to sixteen quarters after becoming known.

The middle plot shows the same IRFs scaled appropriately.

The right plot shows the sum of the scaled IRFs shown in the central figure, where the red line gives the ZLB's location, relative to steady state.

Figure 7 illustrates the potential consequences of this multiplicity.[27] It shows two solutions following a 10 standard deviation demand shock (i.e. a positive shock to β_t). For purely illustrative purposes, we also include a consumption equivalent measure of welfare. This is the quantity Z_t which solves:

$$\mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=0}^{s} \beta_{t+s} \right] \left[\log C_{t+s} - \frac{\psi}{1+\vartheta} L_t^{1+\vartheta} \right] = \mathbb{E}_t \sum_{s=0}^{\infty} \left[\prod_{k=0}^{s} \beta_{t+s} \right] \left[\log(\tilde{C}_{t+s} Z_t) - \frac{\psi}{1+\vartheta} \tilde{L}_t^{1+\vartheta} \right],$$

where $\tilde{C}_t$ and $\tilde{L}_t$ are the values consumption and labour supply would take were prices flexible. Z_t will be less than one in steady-state due to the distortion of price-dispersion. However, these welfare calculations come with two caveats. Firstly, all our calculations here are under perfect foresight, so our welfare measure is not capturing any of the effects of uncertainty. Secondly,

[27] This figure, like the previous ones in this subsection, may be generated by running "GeneratePlots" within the "Examples/FernandezVillaverdeEtAl2015" directory of DynareOBC.

our welfare measure is based on an underlying first order approximation, which is likely to be unreliable given such big shocks. To mitigate this, we calculate welfare and other variables in a way which introduces no further error beyond the approximation error coming from the four endogenous variables, inflation, labour supply, price dispersion and the firms' auxiliary variable. Thus, all equations except the four marked with $(*)$ will hold exactly, e.g. it will always be true that $C_t + G_t = Y_t = \frac{A_t}{\nu_t} L_t$ ensuring that consumption levels are feasible given labour supply and price dispersion. Despite this, approximation error is likely to be substantial. With these caveats in mind, we see that while welfare actually improves in the "fundamental" solution (due to the reduction in price dispersion), in the second solution consumption equivalent welfare falls by about 12%.

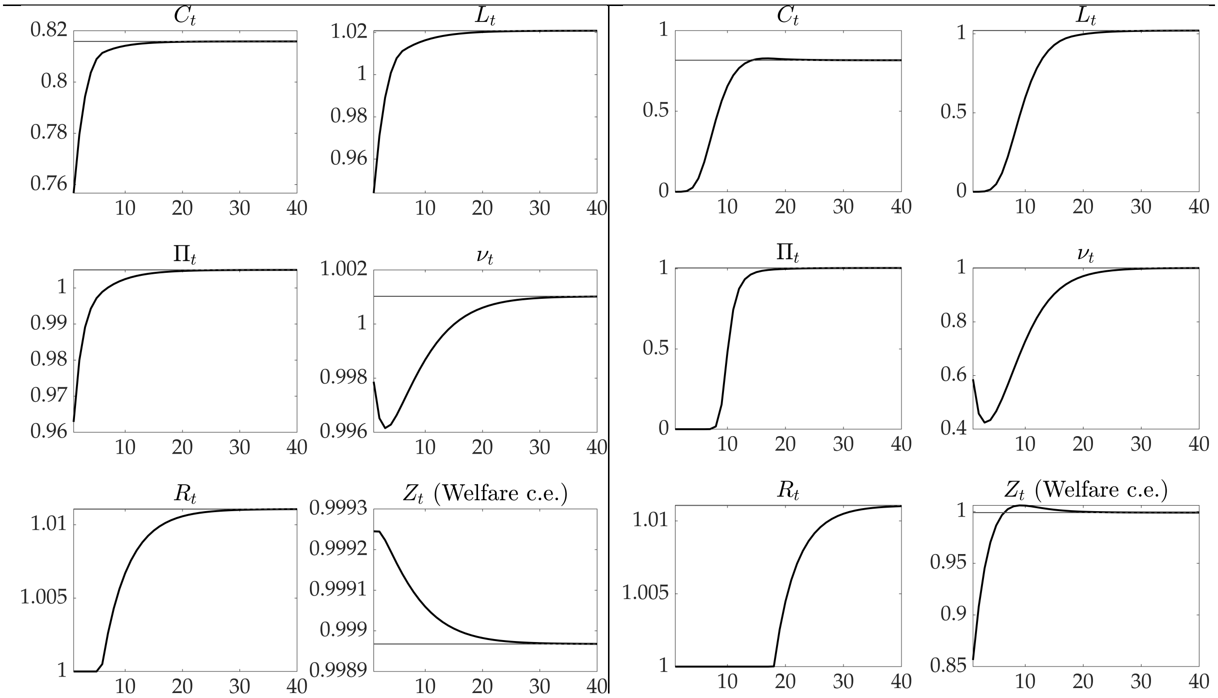


Figure 7: A "good" solution (left 6 panels) and a "bad" solution (right 6 panels), following a 10 standard deviation demand shock in the Fernández-Villaverde et al. (2015) model.

All variables are in levels. The calculation of the welfare consumption equivalent is detailed in the text.

The situation is quite different under price level targeting. In particular, if we replace inflation in the monetary rule with the price level relative to its linear trend, which evolves according to:

$$\log P_t = \log P_{t-1} + \log\left(\frac{\Pi_t}{\Pi}\right), \tag{10}$$

then with $T = 200$, the lower bound from Proposition 2 implies that $\varsigma > 0.003$, and hence that the model is always feasible, in the sense of Definition 6 (Feasibility). Furthermore, even with $T = 1000$, M is a P-matrix by our sufficient conditions from Corollary 7.[28] This is strongly suggestive of uniqueness even for arbitrarily large T, given the reasonably short-lived dynamics of the model.

Appendix E.3: The Smets & Wouters (2003; 2007) models, and the Adjemian, Darracq Pariès & Moyen (2007) model

Smets & Wouters (2003) and Smets & Wouters (2007) are prototypical medium-scale linear DSGE models, featuring assorted shocks, habits, price and wage indexation, capital (with adjustment costs and variable utilisation) and 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, they are broadly similar models, and any differences in their behaviour chiefly stems from differences in the estimated parameters. Since both models are 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 in the presence of the ZLB, we augment each model with a ZLB on nominal interest rates, and evaluate the properties of each model's M matrix at the estimated posterior-modes from the original papers. To minimise the deviation from the original papers, we do not introduce an auxiliary variable for shadow nominal interest rates, so the monetary rules take the form of $i_t = \max\{0, \rho_i i_{t-1} + (1 - \rho_i)(\cdots) + \cdots\}$, in both cases. Our results would be similar with a shadow nominal interest rate.

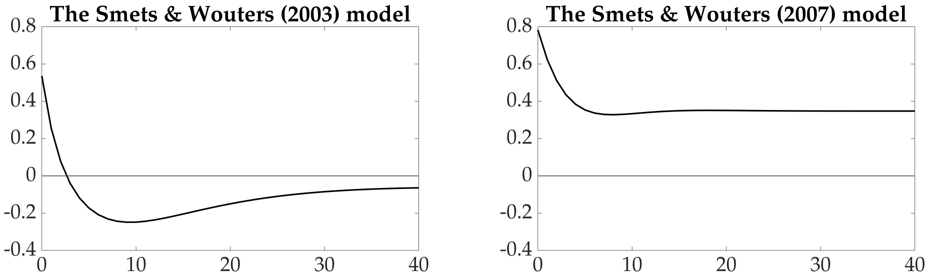
Recall that for ZLB models, the diagonal of the M matrix captures whether positive news shocks to monetary policy raise nominal interest rates in the period in which the shock hits. If this diagonal ever goes negative, then the M matrix cannot be a P-matrix, and hence the model will have multiple solutions in at least some states. In Figure 1,[29] we plot the diagonal of the M

[28] These results are among those that may be generated by running "GenerateDeterminacyResults" within the "Examples/FernandezVillaverdeEtAl2015" directory of DynareOBC.

[29] Details on replicating all of the results in this subsection are given below.

matrix for the two models. 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 & Wouters (2003) model, there is not always a unique equilibrium.

Figure 8: The diagonals of the M matrices for the Smets & Wouters (2003; 2007) models



The horizontal axis gives the index into the M matrix, t. The vertical axis gives the value of $M_{t,t}$.

For the Smets & Wouters (2007) model, numerical calculations reveal that for $T < 9$, M is a P-matrix. However, with $T \geq 9$, the top-left 9×9 sub-matrix of M has negative determinant, so for $T \geq 9$, M is not a P-matrix.[30] Thus, this model also has multiple equilibria. While placing a larger coefficient on inflation in the Taylor rule can make the Euro area picture more like the US one, with a positive diagonal to the M matrix, even with incredibly large coefficients, M remains a non-P-matrix for both models. This is driven by the real and nominal rigidities in the model reducing the average value of the impulse response to a positive news shock to the monetary rule. Following such a shock's arrival, the rigidities help ensure that the fall in output is persistent. Prior to its arrival, consumption habits and capital/investment adjustment costs help produce a larger anticipatory recession. 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. Below we given an example of multiplicity in the Smets & Wouters (2003) model, and discusses the economic relevance of such multiplicity.

In addition, it turns out that for neither model is M an S-matrix even with $T = 1000$, and thus for both models there are some initial states (possibly augmented with monetary policy news shocks) for which no solution exists which escapes the bound after at most 250 years.

[30] The value of T at which the model switches from M being a P-matrix to M not being a P-matrix is parameter dependent. It reflects the strength of the model's endogenous persistence.

This is strongly suggestive of non-existence for some initial states even for arbitrarily large T. This is reinforced by the fact that for the Smets & Wouters (2007) model, with $T = 1000$, Proposition 2 gives that $\varsigma \le 0 +$ numerical error (with ς as defined in Subsection 4.4), which is suggestive of non-existence even for infinite T.

With a response to the price level, the situation is very different. Suppose that in both models we replace the monetary rule by a simple rule responding to the price level and output growth, so it becomes:

$$i_t = \max\left\{0, \rho_i i_{t-1} + (1-\rho_i)\log\left(P_t \frac{Y_t}{Y_{t-1}}\right)\right\},$$

where ρ_i is as in the original model, Y_t is real GDP and where the price level P_t evolves according to $\log P_t = \log P_{t-1} + \log\left(\frac{\Pi_t}{\Pi}\right)$. Then, with $T = 1000$, for either model the sufficient conditions we introduce in Appendix C.1 imply that M is a P-matrix. Hence, the models have a unique solution conditional on escaping after at most 250 years. Additionally, we have that $\varsigma > 0.036$ for the Euro area model with this monetary rule, and that $\varsigma > 0.009$ for the US one (with ς as defined in Subsection 4.4). Hence, Corollary 4 implies that the model always has a feasible path. This is a necessary condition for existence of a solution for any initial state. As one would expect, these results are also robust to departures from equal, unit, coefficients on prices and output growth. Thus, price level targeting again appears to be sufficient for determinacy in the presence of the ZLB.

The intuition again comes down to the sign of the response to monetary policy (news) shocks. With the price level in the Taylor rule, the reduction in prices brought about by a positive monetary policy (news) shock must be followed eventually by a counter-balancing increase. But if inflation is higher in future, then real rates are lower today, meaning that consumption, output, inflation and nominal rates will all be relatively higher today. This ensures that positive monetary policy (news) shocks have sufficiently positive effects on nominal rates to prevent self-fulfilling jumps to the bound. Thus, in the presence of the ZLB, a positive response to the price level is the equivalent of the Taylor principle.

The plots above and the related results from the main paper may be generated by running "GeneratePlots" within both the "Examples/SmetsWouters2003" and

"Examples/SmetsWouters2007" directories of DynareOBC. The determinacy results may be generated by running:

"GenerateDeterminacyResults"

within both the:

"Examples/SmetsWouters2003" and "Examples/SmetsWouters2007"

directories of DynareOBC.

Plausibility of multiplicity at the ZLB

We need to answer two key questions to establish the economic relevance of self-fulfilling spells at the ZLB. Firstly, is the coordination of beliefs needed to sustain the equilibrium plausible? Secondly, do such equilibria feature reasonable movements in macroeconomic variables? It is true that self-fulfilling jumps to the ZLB may feature implausibly large falls in output and inflation. This is closely related to the so-called "forward guidance puzzle" (Carlstrom, Fuerst & Paustian 2015; Del Negro, Giannoni & Patterson 2015).[31] However, if interest rates are already low (due to a recession), then a much smaller self-fulfilling "news" shock is needed to produce a jump to the ZLB. Thus, there will be a much more moderate drop in output and inflation. Furthermore, with interest rates low, it takes a smaller movement in confidence for people to expect to hit the ZLB. Even more plausibly, if the economy is already at the ZLB, then small changes in confidence could easily select an equilibrium featuring a longer spell there than in the equilibrium that leaves fastest. Indeed, there is no good reason people should coordinate on the equilibrium with the shortest time at the ZLB.

[31] McKay, Nakamura & Steinsson (2017) point out that these implausibly large responses to news are muted in models with heterogeneous agents, and give a simple "discounted Euler" approximation that produces similar results to a full heterogeneous agent model. While including a discounted Euler equation makes it harder to generate multiplicity (e.g. reducing the parameter space with multiplicity in the Brendon, Paustian & Yates (2013) model), when there is multiplicity, the resulting responses are much larger, as the weaker response to news means the required endogenous "news" needs to be much greater in order to drive the model to the bound.

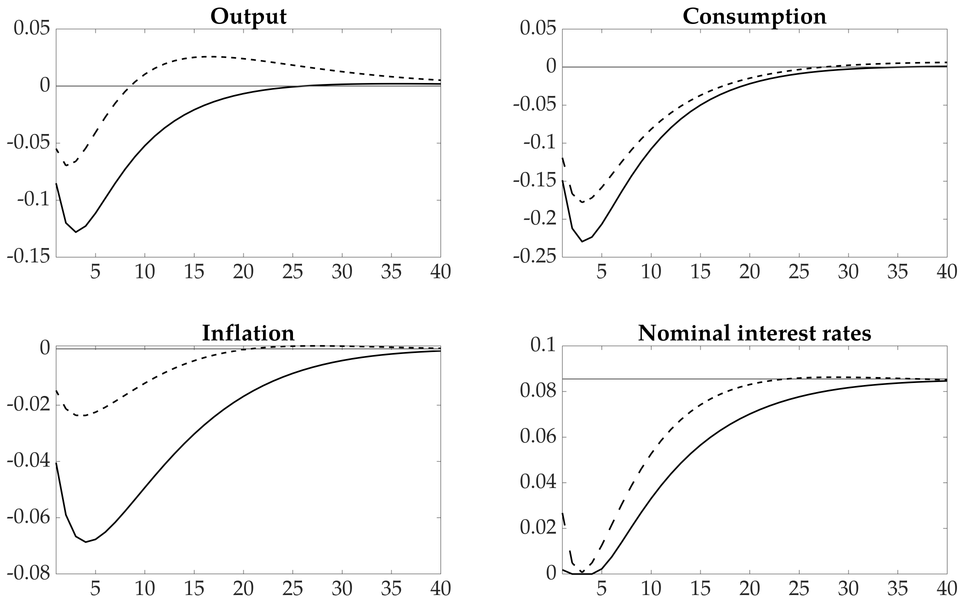


Figure 9: Two solutions following a preference shock in the Smets & Wouters (2003) model.

All variables are in logarithms. Inflation and nominal interest rates are annualized.

The dashed line is a solution which does not hit the bound.

The solid line is an alternative solution which does hit the bound.

As an illustration, in Figure 9 we plot the impulse response to a large magnitude preference shock (scaling felicity), in the Smets & Wouters (2003) model.[32] The shock is not quite large enough to send the economy to the ZLB[33] in the standard solution, shown with a dashed line. However, there is an alternative solution in which the economy jumps to the bound one period after the initial shock, remaining there for three periods. (This is the solution featuring the shortest possible positive time at the ZLB.) While the alternative solution features larger drops in output and inflation, the falls are not much larger than the magnitude of the crisis, with Eurozone GDP and consumption falling about 20% below a pre-crisis log-linear trend, and the largest drop in annualized Eurozone consumption inflation from 2008q3 to 2008q4 being

[32] The shock is 22.5 standard deviations. While this is implausibly large, the economy could be driven to the bound with a run of smaller shocks. It is also worth recalling that the model was estimated on the great moderation period, so the estimated standard deviations may be too low, and the real interest rate too high. Finally, recent evidence (Cúrdia, Del Negro & Greenwald 2014) suggests that the shocks in DSGE models should be fat tailed, making large shocks more likely.

[33] Since the Smets & Wouters (2003) model does not include trend growth, it is impossible to produce a steady state value for nominal interest rates that is consistent with both the model and the data. We choose to follow the data, setting the steady state of nominal interest rates to its mean level over the same sample period used by Smets & Wouters (2003), using data from the same source (Fagan, Henry & Mestre 2005).

around 4.4%.[34] Of course, we would not want to push this as the sole explanation for the depth of the great recession in the Eurozone, but it is possible that multiplicity of equilibria may have played a role.

The Adjemian, Darracq Pariès & Moyen (2007) model

We can get a sense of the potential welfare benefits of a switch to price level targeting by comparing equilibria with and without a response to the price level in a closely related model, that of Adjemian, Darracq Pariès & Moyen (2007).[35] This is essentially a re-estimated version of the Smets & Wouters (2003) model.[36] It is convenient for our purposes though because whereas the original Smets & Wouters (2003) model was hand-linearized, with some ad hoc changes made only to the linearized equations, the Adjemian, Darracq Pariès & Moyen (2007) model is presented in its fully non-linear form, and welfare measures are derived. The measure of consumption equivalent welfare we use here is much as before. It is the amount of extra consumption services flow you would have to give to an inhabitant of the flexible price version of the model to make them indifferent between their economy and that of the model.[37] Unlike in the Fernández-Villaverde et al. (2015) model though, here it is assumed that non-updating prices are indexed to a combination of lagged inflation, and the steady-state level of inflation. Thus, there is no price-dispersion in steady-state, so steady-state welfare equals that of the flexible price economy.

[34] Data was again from the area-wide model database (Fagan, Henry & Mestre 2005).

[35] An implementation of this model is contained within DynareOBC in the file "Examples/AdjemianDarracqPariesMoyen2007/SWNLWCD.mod".

[36] The only significant difference is that habits are internal, not external.

[37] Habits slightly complicate this. Following Adjemian, Darracq Pariès & Moyen (2007), we assume that it is the habit adjusted consumption flow that is adjusted in the flexible price economy to derive the consumption equivalent welfare. I.e. $(C_{t+s} - hC_{t+s-1})$ in the utility function is replaced with $(C_{t+s} - hC_{t+s-1})Z_t$, where Z_t captures the consumption equivalent welfare.

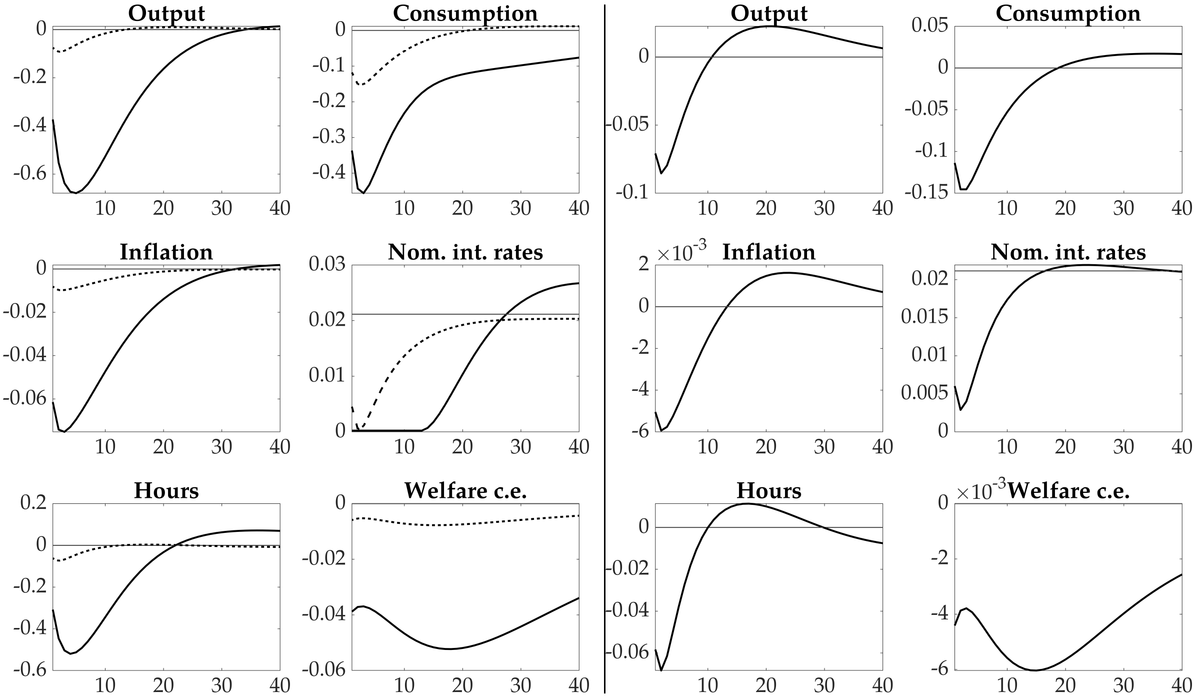


Figure 10: Two solutions following a preference shock in the Adjemian, Darracq Pariès & Moyen (2007) model, without (left 6 panels) and with (right 6 panels) a response to the price level.

All variables are in logarithms. The calculation of the welfare consumption equivalent is detailed in the text.

The dashed line is a solution which does not hit the bound.

The solid line is an alternative solution which does hit the bound in the absence of price level targeting.

The two solutions are identical with a response to the price level.

Our impulse response exercise in Figure 10 follows that of Figure 9 above,[38] and without a response to the price level, the responses of other variables are qualitatively very similar to those in that figure.[39] However, suppose we add a response to the price level to the monetary rule, so it becomes:

$$i_t = \max\{0, \rho_i i_{t-1} + (1-\rho_i)\log(P_t) + \text{other terms from the original model}\},$$

where ρ_i is as in the original model, and where the price level P_t again evolves per equation (10). Then the second solution no longer exists, so the welfare outcome is much improved (a

[38] This figure may be generated by running "RunExample" within the "Examples/AdjemianDarracqPariesMoyen2007" directory of DynareOBC.

[39] In this case, we need a slightly larger shock for a comparable exercise. It is now 24.5 standard deviations rather than 22.5 standard deviations used in Figure 9.

0.6% drop rather than a 5% drop). As in the Fernández-Villaverde et al. (2015) example though, this is again subject to the same caveats on accuracy.[40]

Appendix E.4: The deflationary steady state

We now analyse a version of the model of Subsection 2.1 and 2.5 from the paper in which agents expect the economy to return to the deflationary steady state. We allow for the possibility of a response to a geometrically weighted average of inflation, $\tilde{p}_t$. The model's equations are:

$$r_t + \pi_{t+1} = i_t = \max\{0, r_t + \phi\pi_t + \chi\tilde{p}_t\}$$

$$\tilde{p}_t = \rho\tilde{p}_{t-1} + \pi_t$$

As $\rho \to 1$ this approaches a model with a price level target. Working with this average inflation targeting version ensures that all variables have a finite steady state, even when $i = 0$ in steady state.

We define $s_t := r_t + \phi\pi_t + \chi\tilde{p}_t$, and $\tilde{\imath}_t := \max\{-s_t, 0\}$, so $i_t = s_t + \tilde{\imath}_t$. Note that $\tilde{\imath}_t$ is positive in the deflationary steady state, so this transformed OBC does not bind in the deflationary steady state. In the vicinity of the deflationary steady state, the model is indeterminate. The approach of Farmer, Khramov & Nicolò (2015) suggests introducing the new equation $\pi_t = e_{t-1} + \eta_t$, where the new variable $e_t = \mathbb{E}_t\pi_{t+1}$ gives inflation expectations and where η_t is a "sunspot" shock (since anticipated sunspot shocks make little sense, we assume $\eta_t = 0$ for $t \neq 0$). Then the complete model is:

$$\pi_t = e_{t-1} + \eta_t,$$

$$r_t + e_t = i_t = s_t + \tilde{\imath}_t,$$

$$\tilde{p}_t = \rho\tilde{p}_{t-1} + e_{t-1} + \eta_t,$$

$$s_t = r_t + \phi e_{t-1} + \phi\eta_t + \chi\tilde{p}_t,$$

$$\tilde{\imath}_t = \max\{-s_t, 0\}.$$

This is a completely backward-looking model, so the results of Corollary 2 apply. Thus, in order to test whether or not there is uniqueness, we just have to look at the impact of a shock to

[40] While the economy is moving less far from its steady state following this shock than in the Fernández-Villaverde et al. (2015) example, here all variables, including welfare, are in first order approximations.

the bounded equation, ignoring the bound. This requires us to solve the model with extra shock ν_t (but without the r_t or η_t shocks) given by:

$$\pi_t = e_{t-1},$$
$$r + e_t = i_t = s_t + \tilde{\imath}_t,$$
$$\tilde{p}_t = \rho\tilde{p}_{t-1} + e_{t-1},$$
$$s_t = r + \phi e_{t-1} + \chi\tilde{p}_t,$$
$$\tilde{\imath}_t = -s_t + \nu_t.$$

The solution has:

$$e_t = -r + \nu_t,$$
$$s_t = r + (\phi + \chi)e_{t-1} + \chi\rho\tilde{p}_{t-1},$$
$$\tilde{\imath}_t = -r - (\phi + \chi)e_{t-1} - \chi\rho\tilde{p}_{t-1} + \nu_t.$$

Thus, positive shocks to the bounded equation increase the bounded variable one for one, i.e., $M_{1,1} = 1 > 0$. This gives uniqueness, conditional on existence, via the following general result:

Corollary 8 Consider a purely backward-looking otherwise linear model with an OBC. Then:

1) If $M_{1,1} > 0$, then for any x_0 and $(\varepsilon_t)_{t=1}^{\infty}$, if there exists a path satisfying the model's equations and eventually escaping the bound, then that path is unique.[41]

Furthermore, suppose the model has at least one t-dated shock with a non-zero impact on i_t (if the model has a shock to the bounded equation, then $M_{1,1} \neq 0$ is sufficient for this), then:

2) If $M_{1,1} \leq 0$ then for any $T > 0$, there exists x_0 and $(\varepsilon_t)_{t=1}^{\infty}$ with $\varepsilon_t = 0$ for $t > T$ such that there are multiple paths $(x_t)_{t=1}^{\infty}$ satisfying the model's equations from period 1 to T and satisfying the model's equations without the OBC (i.e. with the max removed) from period $T + 1$ onwards.

Appendix F: Relationship between multiplicity under perfect-foresight, and multiplicity under rational expectations

Our results here will apply to any non-linear dynamic model, not just otherwise linear models with occasionally binding constraints. Before starting, we introduce a little notation that

[41] Existence of a path escaping the bound is not guaranteed in general. Consider the model $z_t = \frac{1}{2}z_{t-1} - \frac{4}{5}\left(i_t - \frac{4}{5}z_t\right)$, $i_t = \max\{-1, \frac{4}{5}z_t\}$, with $x_t' = [i_t, z_t]$. If $z_0 = -\frac{40}{7}$, then $z_1 = -\frac{40}{7}$ is the unique solution, despite this having $M_{11} = \frac{9}{25} > 0$.

will be helpful in the below. For any subsets of Euclidean space, A and B, and any function $h: A^J \to B$, we define $\|h\|_\infty := \sup_{a_1,\dots,a_J \in A} \|h(a_1, \dots, a_J)\|_2$, and $L(h) := \sup_{a_1,\dots,a_J,a_1^*,\dots,a_J^* \in A_J} \frac{\|h(a_1,\dots,a_J)-h(a_1^*,\dots,a_J^*)\|_2}{\max\{\|a_1-a_1^*\|_2,\dots,\|a_J-a_J^*\|_2\}}$, so $L(h)$ is the Lipschitz modulus of h. We extend this notation to subscripted functions by also taking suprema over the subscripts. So, for example if $h \in \prod_{j=1}^{J}(A_j \to B)$, then $\|h\|_\infty = \sup_{j=1,\dots,J} \|h_j\|_\infty$ and $L(h) = \sup_{j=1,\dots,J} L(h_j)$.

Let x_t be a vector of the model's endogenous variables, with $x_t \in \mathcal{X} \subseteq \mathbb{R}^n$. Similarly, let ε_t be a vector of the model's exogenous i.i.d. shocks, with $\varepsilon_t \in \mathcal{E} \subseteq \mathbb{R}^m$, where $0 \in \mathcal{E}$. We assume that with probability $1-\sigma$, $\varepsilon_t = 0$, while with probability σ, ε_t is drawn from a probability distribution over $\mathcal{E}$ with measure p. This distribution may be either continuous or discrete. Thus, $\sigma = 0$ corresponds to the perfect foresight case, while when $\sigma = 1$, the distribution of ε_t is unrestricted. We assume t dated variables are known at t. There is no requirement that either x_t or ε_t be in any sense "minimal". For example, ε_t may contain non-fundamental shocks with no impact on the value of the model's equations, except perhaps through beliefs.

We assume that at any point in time, the economy can be in any one of a set K of "regimes". Both the policy functions, and the model's equations may differ across these regimes. Thus, these regimes can capture both switching sunspot solutions (with differing policy functions but identical model equations) and switching model properties (with the model equations switching). If the model equations do not vary over K, then in the limit as uncertainty disappears, these regimes will capture $|K|$ different perfect foresight solutions to the model. (K may be finite or countably infinite.) We denote the regime in period t by k_t. Within each regime, the policy functions and model equations may be a function of the length of time the economy has been in the current regime, denoted by s_t. $s_t = 1$ in the first period in a new regime, $s_t = 2$ in the second, and so on.

At the start of each period a binary "transition shock" is realised. With probability $1-\delta$, the transition shock does not hit, and the economy will remain in the regime it was in last period. However, with probability δ, the economy is hit with the transition shock, and transitions to another regime according to the period t Markov transition matrix $\Omega_t := \left[\omega_{k,l}^{(t)}\right]_{k,l \in K}$. $\omega_{k,l}^{(t)} \in [0,1]$ gives the probability of transitioning from regime k to regime l at the start of period t, conditional on the transition shock hitting. Rows of Ω_t sum to 1. If $\omega_{k,k}^{(t)} \neq 0$ for some k, then if $k_{t-1} = k$, there is a $\delta\omega_{k,k}^{(t)}$ chance of remaining in regime k at t but with the "clock" reset, as if the economy had just arrived at regime k. We assume that for all $t \in \mathbb{Z}$, $k,l \in K$, $\omega_{k,l}^{(t)} = \omega_{k,l,s_t,\varepsilon_t}(x_{t-1})$ where $\omega_{k,l,s,e}: \mathcal{X} \to [0,1]$ is a Lipschitz continuous function with $L(\omega_{k,l,s,e}) < \infty$ for all $k,l \in K$, $s \in \mathbb{N}^+$ and $e \in \mathcal{E}$. This allows transition probabilities to be deterministic functions of the current state and shock. We further assume that:

$$L_\Sigma(\omega) := \sup_{k\in K, s\in\mathbb{N}^+, e\in\mathcal{E}} \sum_{l\in K} L(\omega_{k,l,s,e}) < \infty.$$

We assume that the model's equations (first order conditions, laws of motion, etc.) are in the general form:

$$0 = \mathbb{E}_t \mathcal{f}_{k_t,s_t,\varepsilon_t}(x_{t-1}, x_t, x_{t+1}),$$

where $\mathcal{f}_{k,s,e}: \mathcal{X}^3 \to \mathbb{R}^n$ is a Lipschitz continuous function with $L(\mathcal{f}_{k,s,e}) \leq L(\mathcal{f}) < \infty$ for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$. We impose no stability requirement beyond $x_t \in \mathcal{X}$. The rational expectations solutions we find will be near to a corresponding perfect foresight one, so by limiting the perfect foresight equilibria considered, we can rule out explosive equilibria. Such equilibria could also be ruled out by bounding $\mathcal{X}$.

The restriction to Lipschitz continuous $\mathcal{f}$ is almost without loss of generality. Note that all otherwise linear models with occasionally binding constraints are Lipschitz continuous. More generally, virtually all economic models will result in a Lipschitz $\mathcal{f}$, perhaps after some variable transformations such as using the logarithm not the level of the capital stock as an endogenous variable. Lipschitz continuity just rules out functions that are not differentiable almost everywhere, or for which the derivative is not almost everywhere bounded.

Given some σ and δ, we write $g_{k,s,e}^{(\sigma,\delta)}: \mathcal{D}_{k,s} \to \mathcal{X}$ for the (unknown) policy function in the s^{th} period in regime k with shock e, meaning that for all t, $x_t = g_{k_t,s_t,\varepsilon_t}^{(\sigma,\delta)}(x_{t-1})$. $\mathcal{D}_{k,s} \subseteq \mathcal{X}$ is the x-domain of definition of the policy functions, taken to be independent of σ and δ. This may be less than then entire space due to non-existence in some areas. We write $g_{k,s,e}^* := g_{k,s,e}^{(0,0)}$ for the known policy function in the perfect foresight case with $\sigma = 0$ and $\delta = 0$.

Let $\mathcal{B}$ denote the closed unit ball in $\mathbb{R}^n$. We assume that there exists $\zeta > 0$ such that $g_{k,s,e}^*(\mathcal{D}_{k,s}) + \zeta\mathcal{B} \subseteq \mathcal{D}_{k,s+1} \cap \mathcal{D}_{l,1} \subseteq \mathcal{X}$ for all $k,l \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$. This ensures that under perfect foresight, x_t is guaranteed to stay within the domain of definition of the policy function in all periods, even if a probability zero transition shock hit. The addition of $\zeta\mathcal{B}$ ensures that small perturbations of the images will remain inside the relevant domains.

Our goal is to establish existence of the policy function for some $\sigma > 0$ and $\delta > 0$. We look for policy functions that have $\left\|g^{(\sigma,\delta)} - g^*\right\|_\infty \le \zeta$, meaning they remain close enough to the perfect foresight policy function for the domains to remain valid. We also assume $L\left(g^{(\sigma,\delta)}\right) < \infty$. This is reasonable given the Lipschitz continuity of f. For example, otherwise linear models with occasionally binding constraints have Lipschitz policy functions at least over any convex sets on which their solution is unique, by a corresponding result for LCPs (Mangasarian & Shiau 1987). Finally, we require that there exist values $A^{(\sigma,\delta)} > 0$ and $B^{(\sigma,\delta)} > 0$ such that for all $k, l \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s+1}$:

$$\left\|g^{(\sigma,\delta)}_{k,s+1,e}(x) - g^{(\sigma,\delta)}_{k,s+1,0}(x)\right\|_2 \le A^{(\sigma,\delta)},$$
$$\left\|g^{(\sigma,\delta)}_{l,1,e}(x) - g^{(\sigma,\delta)}_{k,s+1,e}(x)\right\|_2 \le B^{(\sigma,\delta)}.$$

This is a substantially weaker condition than requiring $g^{(\sigma,\delta)}$ to be bounded. The first condition states that the policy function with a shock should not be too different to the one without a shock. The second states that the policy functions should not differ too much across regimes. $A^{(\sigma,\delta)}$ and $B^{(\sigma,\delta)}$ may be arbitrarily large, so this still permits large responses to shocks and drastically different behaviour across regimes. While the first condition may require shocks to be bounded in some models, these bounds on shocks can be arbitrarily large.

To be a solution, for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$, these policy functions must satisfy:

$$\begin{aligned}
0 = {} & (1-\delta)(1-\sigma) f_{k,s,e}\left(x, g^{(\sigma,\delta)}_{k,s,e}(x), g^{(\sigma,\delta)}_{k,s+1,0}\left(g^{(\sigma,\delta)}_{k,s,e}(x)\right)\right) \\
& + (1-\delta)\sigma \int_{\mathcal{E}} f_{k,s,e}\left(x, g^{(\sigma,\delta)}_{k,s,e}(x), g^{(\sigma,\delta)}_{k,s+1,\varepsilon}\left(g^{(\sigma,\delta)}_{k,s,e}(x)\right)\right) dp(\varepsilon) \\
& + \delta(1-\sigma) \sum_{l \in K} \omega_{k,l,s,e}(x) f_{k,s,e}\left(x, g^{(\sigma,\delta)}_{k,s,e}(x), g^{(\sigma,\delta)}_{l,1,0}\left(g^{(\sigma,\delta)}_{k,s,e}(x)\right)\right) \\
& + \delta\sigma \sum_{l \in K} \omega_{k,l,s,e}(x) \int_{\mathcal{E}} f_{k,s,e}\left(x, g^{(\sigma,\delta)}_{k,s,e}(x), g^{(\sigma,\delta)}_{l,1,\varepsilon}\left(g^{(\sigma,\delta)}_{k,s,e}(x)\right)\right) dp(\varepsilon).
\end{aligned} \tag{11}$$

This equation just encodes the rules for transitioning between regimes already discussed.

When $\sigma = 0$ and $\delta = 0$, all future uncertainty disappears, and we are left with perfect foresight solutions. Setting $\sigma = 0$ and $\delta = 0$ in equation (11) gives:

$$0 = f_{k,s,e}\left(x, g^*_{k,s,e}(x), g^*_{k,s+1,0}\left(g^*_{k,s,e}(x)\right)\right). \tag{12}$$

These are the standard equations defining perfect foresight policy functions. In this case, the regime never changes from its initial value, and so "clock time", s, gives actual time, t. The perfect foresight iteration $x_t = g^*_{k,t,\varepsilon_1 \mathbb{1}[t=1]}(x_{t-1})$ may converge to a different steady state in different regimes, or for different initial states x_0 and first period shocks ε_1. It may also cycle rather than converging.

We assume that the perfect foresight policy function $g^* = \left(g^*_{k,s,e}\right)_{k\in K, s\in\mathbb{N}^+, e\in\mathcal{E}}$ is known, and satisfies the properties given above. More than this though, we assume that there exists $\lambda > 0$ such that for all sequences $\eta := (\eta_s)_{s\in\mathbb{N}^+} \subseteq \mathbb{R}^n$ with $\|\eta\|_\infty \le \lambda$, we know a modified perfect foresight policy function $g^{*(\eta)}$, where:

$$\eta_1 = f_{k,s,e}\left(x, g^{*(\eta)}_{k,s,e}(x), g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right),$$

for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$, where R is the shift operator defined by $(R\eta)_s = \eta_{s+1}$ for all $s \in \mathbb{N}^+$. We assume $g^{*(0)} = g^*$. This captures the perfect foresight solution to an augmented model in which an additional exogenous "forcing process" is added to each equation. If we can solve the original perfect foresight system, we ought to be able to solve this augmented system for sufficiently small forcing processes. This is very similar to our approach for models with occasionally binding constraints. There we added forcing processes to the bounded equation and showed that the model could still be solved.

For $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$, $x \in \mathcal{D}_{k,s}$ and $(\eta_s)_{s\in\mathbb{N}^+} \subseteq \mathbb{R}^n$ with $\|\eta\|_\infty \le \lambda$, we write $\mathcal{P}^{(\eta)}_{k,s,e}(x)$ for the path followed by the endogenous variables under perfect foresight, when started with the given state and shock, with the given forcing process. Thus:

$$\mathcal{P}^{(\eta)}_{k,s,e}(x) = \left(g^{*(\eta)}_{k,s,e}(x), g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right), g^{*(R^2\eta)}_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right), \dots\right).$$

We assume there exists $C > 0$ such that for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$, $x \in \mathcal{D}_{k,s}$ and $(\eta_{1,s})_{s\in\mathbb{N}^+}, (\eta_{2,s})_{s\in\mathbb{N}^+} \subseteq \mathbb{R}^n$ with $\|\eta_1\|_\infty \le \lambda$ and $\|\eta_2\|_\infty \le \lambda$:

$$\left\|\mathcal{P}^{(\eta_1)}_{k,s,e}(x) - \mathcal{P}^{(\eta_2)}_{k,s,e}(x)\right\|_\infty \le C\|\eta_1 - \eta_2\|_\infty,$$

and that there exists $D > 0$ such that for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$, $x_1, x_2 \in \mathcal{D}_{k,s}$ and $(\eta_s)_{s\in\mathbb{N}^+} \subseteq \mathbb{R}^n$ with $\|\eta\|_\infty \le \lambda$:

$$\left\|\mathcal{P}^{(\eta)}_{k,s,e}(x_1) - \mathcal{P}^{(\eta)}_{k,s,e}(x_2)\right\|_\infty \le D\|x_1 - x_2\|_2.$$

This means that the future perfect-foresight path of the endogenous variables when started in state (k, s, e, x) is Lipschitz both in the forcing process η and in the initial state x. I.e., small changes in the forcing process or initial state will produce small changes in the future path followed by the endogenous variables. For determinate linear models, the Lipschitz continuity in the initial state is immediate, and the Lipschitz continuity in the forcing process follows from the characterization of the solution under such a forcing process given in Appendix H.4. For otherwise linear models with occasionally binding constraints, this follows from the corresponding result for LCPs (Mangasarian & Shiau 1987) providing $\mathcal{D}_{k,s}$ and λ are chosen to ensure there is always a unique solution. We give three simple examples of checking these forcing process conditions in the next subsection.

Theorem 3 (Restated) Under the conditions outlined in the text above (from Appendix F), there exists $\gamma > 0$ and $\xi \in (0,1)$ such that for all $\sigma < \xi$ and $\delta < \xi$, there exists a policy function $g^{(\sigma,\delta)} = \left(g_{k,s,e}^{(\sigma,\delta)}\right)_{k\in K, s\in\mathbb{N}^+, e\in\mathcal{E}}$ that solves the model (equation (11)), and satisfies the other policy function conditions given above. Moreover, $\left\|g^{(\sigma,\delta)} - g^*\right\|_\infty \leq \gamma \max\{|\sigma|, |\delta|\}$.

We prove this in Appendix H.6. Note that the proof is constructive, so this could form the basis of an effective algorithm for computing global solutions to non-linear rational expectations models. Theorem 3 (Restated) is a powerful tool for proving the existence of rational expectations equilibria for general non-linear models. It implies that if there are multiple discrete solutions under perfect foresight (so $|K| > 1$), then there are a continuum of solutions under rational expectations. Even if $|K| = 1$, then there are still a continuum of solutions under rational expectations if the one perfect-foresight solution is not time invariant, as in the example from Subsection 2.6 of the main paper.

Appendix F.1: Initial simple examples

For a first example of checking the forcing process conditions, consider the minimal Taylor rule model, $0 = i_t - r - 2\pi_t$, $0 = r + \pi_{t+1} - i_t$. When augmented with the η forcing process, this becomes $\eta_{1,t} = i_t - r - 2\pi_t$ and $\eta_{2,t} = r + \pi_{t+1} - i_t$. Thus, $\pi_t = -\sum_{s=0}^{\infty} 2^{-s-1}(\eta_{1,t+s} + \eta_{2,t+s})$ and $i_t = r - \sum_{s=0}^{\infty} 2^{-s}(\eta_{1,t+s} + \eta_{2,t+s}) + \eta_{1,t}$. These are functions of the current and future path of the forcing process, as required. Additionally, the first Lipschitz condition is clearly satisfied with $C = \sqrt{2^2 + 3^2} = \sqrt{13}$, and the second is trivially satisfied as there are no state variables.

Next, consider the zero-peg model $0 = i_t$, $0 = r + \pi_{t+1} - i_t$ which has a perfect foresight solution with $\pi_t = -r$ (amongst others). When augmented with the η forcing process, this states that $\eta_{1,t} = i_t$ and $\eta_{2,t} = r + \pi_{t+1} - i_t = r + \pi_{t+1} - \eta_{1,t}$. This would imply that $\pi_t = -r + \eta_{1,t-1} + \eta_{2,t-1}$ for $t > 1$. However, this is not a function of current and future values of the forcing process, so does not meet the requirements. The failure here is due to the indeterminacy of the zero-peg model.

Luckily, there is a simple fix. We can always transform an indeterminate model into a determinate one (possibly augmented with sunspot shocks) following the method of Farmer, Khramov & Nicolò (2015). We can change the model's equations in this way for only certain k since we allow $f_{k,s,e}$ to vary with k. The solution to this equivalent determinate model will still exist even when η is included. Under this transformation, the zero-peg model becomes $0 = i_t$, $0 = r + e_t - i_t$, with the extra equation $0 = \pi_t - e_{t-1}$. Adding shocks $\eta_{1,t}$, $\eta_{2,t}$ and $\eta_{3,t}$ respectively to the left hand sides of these equations produces the solution: $i_t = \eta_{1,t}$, $e_t = -r + \eta_{1,t} + \eta_{2,t}$ and $\pi_t = e_{t-1} + \eta_{3,t}$. This no longer contains lagged values of the forcing process (though of course they enter indirectly via the state variable e_t). It satisfies the first Lipschitz condition with $C = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6}$, and the second with $D = 1$.

For a final example, consider the simple model presented in Subsection 2.2 of the main paper, which we examined under rational expectations in Subsection 2.6 there. The model has $0 = i_t - \max\{0, r + \phi\pi_t - \psi\pi_{t-1}\}$, $0 = r + \pi_{t+1} - i_t$, with $\phi := 2$ and $\psi \in (0,1)$. The perfect foresight solution has $\pi_1 = -\frac{r}{A}$ and $i_1 = 0$, unconditionally on π_0, while $\pi_t = A\pi_{t-1}$ for $t > 1$, with $A^2 = \phi A - \psi$. Adding shocks $\eta_{1,t}$ and $\eta_{2,t}$ to the model means we want to solve $\eta_{1,t} = i_t - \max\{0, r + 2\pi_t - \psi\pi_{t-1}\}$ and $\eta_{2,t} = r + \pi_{t+1} - i_t$. We fix the domain of π_t to the set $[-\frac{r}{2A} - \frac{r}{2A^2}, \infty)$ (restricting $\mathcal{X}$ and $\mathcal{D}_{k,s}$). This covers the desired perfect foresight solution with $\pi_t = -A^{t-2}r$.

We conjecture that for sufficiently small $\lambda > 0$, if $\|\eta\|_\infty \le \lambda$, there is a solution of the form:

$$\pi_1 = \frac{1}{A}\left[-r + \eta_{1,1} + \eta_{2,1} + \sum_{s=0}^{\infty}(2 - A)^{-s-1}(\eta_{1,2+s} + \eta_{2,2+s})\right],$$

$$i_1 = \eta_{1,1},$$

with, for $t > 1$:

$$\pi_t = A\pi_{t-1} - \sum_{s=0}^{\infty}(2-A)^{-s-1}(\eta_{1,t+s} + \eta_{2,t+s}),$$
$$i_t = r + A^2\pi_{t-1} + \eta_{1,t} - 2\sum_{s=0}^{\infty}(2-A)^{-s-1}(\eta_{1,t+s} + \eta_{2,t+s}).$$

It is easy to verify that the second equation ($\eta_{2,t} = r + \pi_{t+1} - i_t$) is always satisfied, no matter the value of λ. We just have to check the first. In period 1, the first equation holds if and only if $r + 2\pi_1 - \psi\pi_0 \leq 0$. Given that $\pi_0 \geq -\frac{r}{2A} - \frac{r}{2A^2}$, straightforward computation shows that it suffices to take $\lambda \leq (1-A)^2\frac{r}{8}$. In period $t > 1$, the first equation holds if and only if $r + 2\pi_t - \psi\pi_{t-1} \geq 0$, as the candidate solution satisfies $i_t = r + 2\pi_t - \psi\pi_{t-1} + \eta_{1,t}$. Given that $\pi_{t-1} \geq -\frac{r}{2A} - \frac{r}{2A^2}$, for this condition we again need $\lambda \leq (1-A)^2\frac{r}{8}$, by similar computations. Thus, with $\lambda := (1-A)^2\frac{r}{8}$, both conditions are satisfied, and our conjectured solution is indeed a solution. Having established the solution, it is now easy to see that the required Lipschitz conditions are satisfied with $D = \sqrt{A^2 + A^4}$ and:

$$C = \max\left\{\sqrt{\left(\frac{2}{A} + \frac{2}{1-A}\right)^2 + 1^2}, \sqrt{\left(\frac{2}{1-A}\right)^2 + \left(\frac{3+A}{1-A}\right)^2}\right\}.$$

Appendix G: Results from and for dynamic programming

Appendix G.1: The linear-quadratic case

Alternative existence and uniqueness results for the infinite T problem can be established via dynamic programming methods, under the assumption that Problem 2 (OBC) 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 of greater interest.

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

Problem 4 (Linear-Quadratic) 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 $\tilde{\Gamma}: \mathbb{R}^n \to \mathbb{P}(\mathbb{R}^n)$ (where $\mathbb{P}$ denotes the power-set operator) by:

$$\tilde{\Gamma}(x) = \left\{z \in \mathbb{R}^n \middle| 0 \leq \Psi^{(0)} + \Psi^{(1)}\begin{bmatrix} x-\mu \\ z-\mu \end{bmatrix}\right\}, \tag{13}$$

for all $x \in \mathbb{R}^n$. *(Note: $\tilde{\Gamma}(x)$ 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:

$$\widetilde{\mathbb{X}} := \{x \in \mathbb{R}^n | \tilde{\Gamma}(x) \neq \emptyset\}, \tag{14}$$

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

$$\tilde{\mathcal{F}}(x, z) = u^{(0)} + u^{(1)} \begin{bmatrix} x - \mu \\ z - \mu \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - \mu \\ z - \mu \end{bmatrix}' \tilde{u}^{(2)} \begin{bmatrix} x - \mu \\ z - \mu \end{bmatrix}, \tag{15}$$

for all $x, z \in \widetilde{\mathbb{X}}$, 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 $x_0 \in \widetilde{\mathbb{X}}$ is given and $\beta \in (0,1)$, and choose $x_1, x_2, \dots$ to maximise:

$$\liminf_{T\to\infty} \sum_{t=1}^{T} \beta^{t-1} \tilde{\mathcal{F}}(x_{t-1}, x_t) \tag{16}$$

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

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

Assumption 4 $\tilde{u}^{(2)}$ is negative-definite.

In Appendix H.10, below, we establish the following (unsurprising) result:

Proposition 14 If either $\widetilde{\mathbb{X}}$ is compact, or, $\tilde{\Gamma}(x)$ is compact valued and $x \in \tilde{\Gamma}(x)$ for all $x \in \widetilde{\mathbb{X}}$, then for all $x_0 \in \widetilde{\mathbb{X}}$, there is a unique path $(x_t)_{t=0}^{\infty}$ which solves Problem 4 (Linear-Quadratic).

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:

$$\sum_{t=1}^{\infty} \beta^{t-1} \left[\tilde{\mathcal{F}}(x_{t-1}, x_t) + \lambda'_{\Psi,t} \left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} \right] \right], \tag{17}$$

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 $t \in \mathbb{N}^+$:

$$0 = u^{(1)}_{\cdot,2} + \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}' \tilde{u}^{(2)}_{\cdot,2} + \lambda'_t \Psi^{(1)}_{\cdot,2} + \beta \left[u^{(1)}_{\cdot,1} + \begin{bmatrix} x_t - \mu \\ x_{t+1} - \mu \end{bmatrix}' \tilde{u}^{(2)}_{\cdot,1} + \lambda'_{t+1} \Psi^{(1)}_{\cdot,1} \right], \tag{18}$$

$$0 \le \Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}, \qquad 0 \le \lambda_t, \qquad 0 = \lambda_t \circ \left[\Psi^{(0)} + \Psi^{(1)} \begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} \right], \tag{19}$$

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 $\bar{\lambda}$ to be the steady state of λ_t, we require:

$$0 = u^{(1)}_{\cdot,2} + \bar{\lambda}'\Psi^{(1)}_{\cdot,2} + \beta\big[u^{(1)}_{\cdot,1} + \bar{\lambda}'\Psi^{(1)}_{\cdot,1}\big], \tag{20}$$

$$0 \le \Psi^{(0)}, \qquad 0 \le \bar{\lambda}, \qquad 0 = \bar{\lambda} \circ \Psi^{(0)}. \tag{21}$$

In Appendix H.11, below, we prove the following result:

Proposition 15 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 (18) and (19), and that as $t \to \infty$, $x_t \to \mu$ and $\lambda_t \to \bar{\lambda}$, where μ and λ satisfy the steady state KKT conditions given in equations (20) and (21). Then $(x_t)_{t=1}^{\infty}$ solves Problem 4 (Linear-Quadratic). If, further, either condition of Proposition 14 is satisfied, then $(x_t)_{t=1}^{\infty}$ is the unique solution to Problem 4 (Linear-Quadratic), and there can be no other solutions to the KKT conditions given in equations (18) and (19) satisfying $x_t \to \mu$ and $\lambda_t \to \bar{\lambda}$ as $t \to \infty$.

Now, it is possible to convert the KKT conditions given in equations (18) and (19) into a problem in the form of the multiple-bound generalisation of Problem 2 (OBC) quite generally. To see this, first note that we may rewrite equation (18) as:

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

Now, $\tilde{u}^{(2)}_{2,2} + \beta u^{(2)}_{1,1}$ is negative definite, hence we may define $\mathcal{V} := \Psi^{(1)}_{\cdot,2}\big[\tilde{u}^{(2)}_{2,2} + \beta\tilde{u}^{(2)}_{1,1}\big]^{-1}$, so:

$$\begin{aligned}&\Psi^{(0)} + \Psi^{(1)}\begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} \\ &= \Psi^{(0)} + \big(\Psi^{(1)}_{\cdot,1} - \mathcal{V}\tilde{u}^{(2)}_{2,1}\big)(x_{t-1} - \mu) - \mathcal{V}\left[u^{(1)\prime}_{\cdot,2} + \beta\big[u^{(1)\prime}_{\cdot,1} + \tilde{u}^{(2)}_{1,2}(x_{t+1} - \mu) + \Psi^{(1)\prime}_{\cdot,1}\lambda_{t+1}\big]\right] \\ &-\Psi^{(1)}_{\cdot,2}\big[\tilde{u}^{(2)}_{2,2} + \beta\tilde{u}^{(2)}_{1,1}\big]^{-1}\Psi^{(1)\prime}_{\cdot,2}\lambda_t.\end{aligned} \tag{22}$$

Moreover, equation (19) implies that if the k^{th} element of $\Psi^{(0)} + \Psi^{(1)}\begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix}$ is positive, then the k^{th} element of λ_t is zero, so:

$$\Psi^{(0)} + \Psi^{(1)}\begin{bmatrix} x_{t-1} - \mu \\ x_t - \mu \end{bmatrix} = \max\{0, z_t\}, \tag{23}$$

where:

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

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

$$\mathcal{W} := -\operatorname{diag}^{-1}\operatorname{diag}\left[\Psi^{(1)}_{\cdot,2}\big[\tilde{u}^{(2)}_{2,2} + \beta\tilde{u}^{(2)}_{1,1}\big]^{-1}\Psi^{(1)\prime}_{\cdot,2}\right],$$

providing this is positive (it is nonnegative 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 diag^{-1} maps vectors to a matrix with the given vector on the diagonal, and zeros elsewhere.

We claim that we may replace equation (19) with equation (23) without changing the model. We have already shown that equation (19) implies equation (23), so we just have to prove the converse. We continue to suppose equation (18) holds, and thus, so too does equation (22)(22). Then, from subtracting equation (22) from equation (23), we have that:

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

Hence, as $\mathcal{W}$ is a positive diagonal matrix, and the right-hand side is nonnegative, $\lambda_t \geq 0$. Furthermore, the kth element of λ_t is non-negative if and only if the kth element of z_t is non-positive (as $\mathcal{W}$ is a 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 (23). Thus equation (19) is satisfied.

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

Proposition 16 Suppose we are given a problem in the form of Problem 4 (Linear-Quadratic). Then, the KKT conditions of that problem may be placed into the form of the multiple-bound generalisation of Problem 2 (OBC). Let (q_{x_0}, M) be the infinite LCP corresponding to this representation, given initial state $x_0 \in \widetilde{\mathbb{X}}$. 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 4 (Linear-Quadratic). If, further, either condition of Proposition 14 is satisfied, then this LCP has a unique solution for all $x_0 \in \widetilde{\mathbb{X}}$, which gives the unique solution to Problem 4 (Linear-Quadratic), and, for sufficiently large T^*, the finite LCP $\left(q_{x_0}^{(T^*)}, M^{(T^*)}\right)$ has a unique solution $y^{(T^*)}$ for all $x_0 \in \widetilde{\mathbb{X}}$, 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 4 (Linear-Quadratic).

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 the next subsection, we derive similar results for models with more general constraints and objective functions. The proof of this proposition also showed an alternative method for converting KKT conditions into equations of the form handled by our methods.

Appendix G.2: The general case

Here we consider non-linear dynamic programming problems with general objective functions. Consider then the following generalisation of Problem 4 (Linear-Quadratic):

Problem 5 (Non-linear) Suppose $\Gamma: \mathbb{R}^n \to \mathbb{P}(\mathbb{R}^n)$ is a given compact, convex valued continuous function. Define $X := \{x \in \mathbb{R}^n | \Gamma(x) \neq \emptyset\}$, 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 $x_1, x_2, \dots$ 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 5 X is compact.

Then, by Theorem 4.8 of Stokey, Lucas, and Prescott (1989), there is a unique solution to Problem 5 (Non-linear) for any x_0. We further assume the following to ensure that there is a natural point to approximate around:[42]

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

Having defined μ, we can let $\tilde{\mathcal{F}}$ be a second order Taylor approximation to $\mathcal{F}$ around μ, which will take the form of equation (15)(15). Assumption 4 will be satisfied for this approximation thanks to the concavity of $\mathcal{F}$. 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 $\{(x, \mathrm{z}) | x \in X, z \in \Gamma(x)\}$ is convex. Since it is also compact, by Assumption 5, 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 (13) and $\tilde{X}$ defined as in equation (14):

1) $\mu \in \tilde{X} \subseteq X$,

[42] If X is convex, then the existence of a fixed point of the policy function is a consequence of Brouwer's Fixed Point Theorem, but there is no reason the fixed point guaranteed by Brouwer's Theorem should be even locally attractive.

2) for all $x \in X$, there exists $\tilde{x} \in \widetilde{X}$ such that $\|x - \tilde{x}\|_2 < \epsilon$,

3) for all $x \in \widetilde{X}$, $\tilde{\Gamma}(x) \subseteq \Gamma(x)$,

4) for all $x \in \widetilde{X}$, 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 5 (Non-linear) (and which satisfies Assumption 5 and Assumption 6). If the graph of Γ is convex, then we can construct a problem in the form of the multiple-bound generalisation of Problem 2 (OBC) 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 $x_0 \in \widetilde{X}$. 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.[43] We first approximate the graph of Γ by a polytope (i.e. n 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 nonpositive 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:

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

[43] This is often not too much of a restriction, since equality constraints may be substituted out.

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

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

where $u^{(2)}$ is the Hessian of $\mathcal{F}$, then the Lagrangian in equation (17) 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 & 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 reduced choice set (as the quadratic terms are positive definite).

Appendix H: Proofs

Appendix H.1: Proof of the results of Subsection 2.5

To recap, the model with "news" shocks is:

$$r + p_{t+1} - p_t = r + \phi(p_t - p_{t-1}) + \chi p_t + \nu_t,$$

with $\chi > 0$ and $\phi > 1$, so:

$$p_{t+1} = (1 + \phi + \chi)p_t - \phi p_{t-1} + \nu_t.$$

We fix $p_0 = 0$.

We look for a solution in the form $p_t = \sum_{j=-\infty}^{\infty} G_j \nu_{t+j}$, where $\nu_t = 0$ for all $t \leq 0$. Substituting in, we have:

$$\sum_{j=-\infty}^{\infty} G_{j-1} \nu_{t+j} = (1 + \phi + \chi) \sum_{j=-\infty}^{\infty} G_j \nu_{t+j} - \phi \sum_{j=-\infty}^{\infty} G_{j+1} \nu_{t+j} + \nu_t,$$

so from matching coefficients, we have:

$$G_{-1} = (1 + \phi + \chi)G_0 - \phi G_1 + 1,$$

$$\forall j \neq 0, \qquad G_{j-1} = (1 + \phi + \chi)G_j - \phi G_{j+1}.$$

We conjecture that $G_j = G_0 \zeta^j$ for $j \geq 0$ and $G_j = G_0 \eta^{-j}$ for $j \leq 0$, for some $G_0 \in \mathbb{R}$ and $\zeta, \eta \in (-1,1)$. Then:

$$\eta = (1 + \phi + \chi) - \phi\zeta + \frac{1}{G_0},$$

$$1 = (1+\phi+\chi)\zeta - \phi\zeta^2,$$
$$\eta^2 = (1+\phi+\chi)\eta - \phi.$$

Thus:

$$\eta = \frac{1+\phi+\chi-\sqrt{(1+\phi+\chi)^2-4\phi}}{2} = 1 - \frac{\chi}{\phi-1} + \mathrm{O}(\chi^2),$$
$$\zeta = \frac{\eta}{\phi} = \frac{1+\phi+\chi-\sqrt{(1+\phi+\chi)^2-4\phi}}{2\phi} = \frac{1}{\phi}\left(1-\frac{\chi}{\phi-1}\right) + \mathrm{O}(\chi^2),$$
$$G_0 = -\frac{1}{\sqrt{(1+\phi+\chi)^2-4\phi}} = -\frac{1}{\phi-1} + \frac{\phi+1}{(\phi-1)^3}\chi + \mathrm{O}(\chi^2),$$

where, here and in the following, the $\mathrm{O}(\chi^2)$ terms are taken as $\chi \to 0$. Note that:

$$(1+\phi+\chi)^2 - 4\phi = (\phi-1)^2 + 2\chi(1+\phi) + \chi^2 > 0$$

providing that $\chi \geq 0$, so this solution is real as required.

Additionally:

$$\eta = \frac{1+\phi+\chi-\sqrt{(1+\phi+\chi)^2-4\phi}}{2} > \frac{1+\phi+\chi-\sqrt{(1+\phi+\chi)^2}}{2} = 0,$$

and for $\chi > 0$:

$$\begin{aligned}\eta &= \frac{1+\phi+\chi-\sqrt{(\phi-1)^2+2\chi(1+\phi)+\chi^2}}{2} \\ &< \frac{1+\phi+\chi-\sqrt{(\phi-1)^2+2\chi(\phi-1)+\chi^2}}{2} = \frac{1+\phi+\chi-\sqrt{(\phi-1+\chi)^2}}{2} = 1,\end{aligned}$$

again as required.

Substituting back in, we have:

$$\begin{aligned}i_t &= r + p_{t+1} - p_t \\ &= r + G_0\left[\sum_{j=-\infty}^{-1}\eta^{-j}\nu_{t+1+j} + \sum_{j=0}^{\infty}\zeta^j\nu_{t+1+j} - \sum_{j=-\infty}^{0}\eta^{-j}\nu_{t+j} - \sum_{j=1}^{\infty}\zeta^j\nu_{t+j}\right] \\ &= r + G_0\left[\sum_{j=-\infty}^{0}\eta^{-j+1}\nu_{t+j} + \sum_{j=1}^{\infty}\zeta^{j-1}\nu_{t+j} - \sum_{j=-\infty}^{0}\eta^{-j}\nu_{t+j} - \sum_{j=1}^{\infty}\zeta^j\nu_{t+j}\right] \\ &= r + G_0\left[(1-\zeta)\sum_{j=1}^{\infty}\zeta^{j-1}\nu_{t+j} - (1-\eta)\sum_{j=-\infty}^{0}\eta^{-j}\nu_{t+j}\right] \\ &= r - \sum_{j=1}^{\infty}\frac{\nu_{t+j}}{\phi^j} + G_0\left[\sum_{j=1}^{\infty}\left((1-\zeta)\zeta^{j-1} + \frac{1}{G_0\phi^j}\right)\nu_{t+j} - (1-\eta)\sum_{j=-\infty}^{0}\eta^{-j}\nu_{t+j}\right] \\ &= r - \sum_{j=1}^{\infty}\frac{\nu_{t+j}}{\phi^j} + \left[\frac{1}{(\phi-1)^2}\left(\sum_{j=-\infty}^{0}\nu_{t+j} + \sum_{j=1}^{\infty}\frac{\nu_{t+j}}{\phi^{j-1}}\right) + \frac{1}{\phi-1}\sum_{j=1}^{\infty}\frac{(j-1)\nu_{t+j}}{\phi^j}\right]\chi + \mathrm{O}(\chi^2).\end{aligned}$$

Since the partial derivative of the term in square brackets here with respect to ν_s is strictly positive for all $s \in \mathbb{N}^+$, at least for small χ, all of the elements of M must be strictly monotonically increasing in χ. Thus, by Jacobi's formula, for any principal sub-matrix W of M with $W \in \mathbb{R}^{S\times S}$ ($S \leq T$), if $\chi = 0$:

$$\frac{d \det W}{d\chi} = \frac{dW_{S,1}}{d\chi}(-1)^{S-1} \det W_{1:(S-1),2:S} = \frac{dW_{S,1}}{d\chi}\prod_{s=1}^{S-1}(-W_{s,s+1}) > 0,$$

as with $\chi = 0$, W must be strictly upper triangular with negative elements in the upper triangle. Thus, as $\det W = 0$ when $\chi = 0$, for any T, there exists $\overline{\chi}_T \in (0, \infty]$ such that for all $\chi \in (0, \overline{\chi}_T)$, M (of size $T \times T$) is a P-matrix. (Recall that M is a P-matrix if and only if all its principal sub-matrices have positive determinants.)

We have suggestive evidence that this may be true for all $\chi > 0$. Previously we showed that, at least for small χ, the elements of M with $\chi > 0$ are strictly greater than the elements of M with $\chi = 0$. In fact, this holds for all $\chi > 0$. To see this, note that given $G_0 < 0$ and $\eta < 1$, from examining the square bracketed term of the penultimate expression for i_t above, we just need that $(1-\zeta)\zeta^{j-1} + \frac{1}{G_0\phi^j} < 0$ for $j \geq 1$. With $j = 1$ this holds as:

$$\begin{aligned}(1-\zeta)\zeta^{1-1} + \frac{1}{G_0\phi^1} &= \frac{2\phi - 1 - \phi - \chi + \sqrt{(1+\phi+\chi)^2 - 4\phi}}{2\phi} - \frac{2\sqrt{(1+\phi+\chi)^2 - 4\phi}}{2\phi} \\ &= \frac{-\chi + (\phi - 1) - \sqrt{(1+\phi+\chi)^2 - 4\phi}}{2\phi} \\ &= \frac{-\chi + (\phi - 1) - \sqrt{(\phi-1)^2 + 2\chi(1+\phi) + \chi^2}}{2\phi} < 0.\end{aligned}$$

So, using the fact that $\zeta = \frac{\eta}{\phi} < \eta < 1$:

$$0 < 1 - \zeta < -\frac{1}{G_0\phi}.$$

Thus as $0 < \frac{\eta}{\phi} = \zeta = \frac{\eta}{\phi} < \frac{1}{\phi}$, in fact for all $j \geq 1$:

$$(1-\zeta)\zeta^{j-1} < -\frac{1}{G_0\phi^j},$$

as required. Thus, for all $\chi > 0$, the elements of M with $\chi > 0$ are strictly greater than the elements of M with $\chi = 0$.

Appendix H.2: Proof of the rational expectations results from Subsection 2.6

A_δ, B_δ and C_δ must solve:

$$r + (1-\delta)[A_\delta(A_\delta\pi_{t-1} + B_\delta) + B_\delta] + \delta C_\delta = i_t = r + \phi(A_\delta\pi_{t-1} + B_\delta) - \psi\pi_{t-1},$$

$$r + (1-\delta)[A_\delta C_\delta + B_\delta] + \delta C = i_t = 0,$$

where the former captures the $\pi_t = A_\delta\pi_{t-1} + B_\delta$ case, and the latter captures the $\pi_t = C_\delta$ one. Collecting terms implies that $(1-\delta)A_\delta^2 = \phi A_\delta - \psi$, so:

$$A_\delta = \frac{1-\sqrt{1-(1-\delta)\psi}}{1-\delta} \in \left(0, \frac{1-\sqrt{\delta}}{1-\delta}\right) \subseteq (0,1),\text{[44]}$$

while $(1-\delta)(1+A_\delta)B_\delta + \delta C_\delta = \phi B_\delta$ and $r + (1-\delta)[A_\delta C_\delta + B_\delta] + \delta C_\delta = 0$, so:

$$B_\delta = -\frac{\delta r}{2\delta + (1-\delta)[1-(1-\delta)A_\delta]A_\delta} < 0,$$
$$C_\delta = -\frac{r[2-(1-\delta)(1+A_\delta)]}{2\delta + (1-\delta)[1-(1-\delta)A_\delta]A_\delta} < 0.$$

Note that as $\delta \to 0$, $A_\delta \to A$ (as defined in Subsection 2.2 of the main paper), $B_\delta \to 0$ and $C_\delta \to -A^{-1}r$.

We just have to check that this solution does not violate the monetary rule. First suppose it were the case that $\pi_t = A_\delta\pi_{t-1} + B_\delta$ for all t. Then π_t would converge to the pseudo-steady state $\bar{\pi}$ solving $\bar{\pi} = A_\delta\bar{\pi} + B_\delta$. Thus:

$$\bar{\pi} - C_\delta = \left[\frac{1-\psi}{1-A_\delta}\right]\frac{r}{2\delta + (1-\delta)[1-(1-\delta)A_\delta]A_\delta} > 0.$$

Since $\bar{\pi} - C_\delta > 0$, and $A_\delta \in (0,1)$, in fact for all t, $\pi_t \in [C_\delta, \bar{\pi})$, assuming π_0 is in this interval. Now note that:

$$C_\delta + \frac{r + 2B_\delta}{(1-\delta)A_\delta^2} = C_\delta - \frac{r + 2C_\delta}{\psi} = \left[\frac{1-\psi}{A_\delta}\right]\frac{r}{2\delta + (1-\delta)[1-(1-\delta)A_\delta]A_\delta} > 0.$$

So, for all t:

$$\pi_{t-1} \geq C_\delta > -\frac{r + 2B_\delta}{(1-\delta)A_\delta^2},$$

and thus when $\pi_t = A_\delta\pi_{t-1} + B_\delta$, $r + \phi\pi_t - \psi\pi_{t-1} = r + (1-\delta)A_\delta^2\pi_{t-1} + 2B_\delta > 0$, as required. Similarly, for all t:

$$\pi_{t-1} \geq C_\delta > \frac{r + 2C_\delta}{\psi},$$

[44] Determinacy requires that $1-\delta$ times the other root is greater than one. This holds as $1 + \sqrt{1-(1-\delta)\psi} > 1$.

and thus when $\pi_t = C_\delta$, $r + \phi\pi_t - \psi\pi_{t-1} = r + 2C_\delta - \psi\pi_{t-1} < 0$, again as required. This means that the monetary rule is satisfied: the economy is away from the ZLB when the δ shock does not hit, and at the ZLB when it does.

Appendix H.3: Proof of the NK result from Subsection 4.3

Let $\varepsilon > 0$. Consider increasing the bottom left element of M by $\pm\varepsilon$. By Sylvester's determinant theorem, this increases the determinant of M by $\pm\varepsilon(-1)^{T+1}\det M_{1:T-1,2:T}$.

Write $a_k = M_{T-k,T}$ for $k \geq 0$. Since M is Toeplitz, the definition of a_k is independent of T, so we may consider a_k defined for all $k \in \mathbb{N}$.

To make the determinant of the resulting matrix negative (for an appropriate choice of $\pm$), we need $\frac{|\det M_{1:T-1,2:T}|}{a_0^T} \geq \frac{1}{\varepsilon}$.

With M upper triangular, $M_{1:T-1,2:T}$ is upper Hessenberg. It is also Toeplitz as M is. Hence, by Theorem 1 of Cahill et al. (2002), $\det M_{1:T-1,2:T} = d_{T-1}$, where $d_0 = 1$ and:

$$d_t = \sum_{k=0}^{t-1} (-a_0)^k a_{k+1} d_{t-k-1},$$

for $t > 0$. Write $e_t = \frac{d_t}{a_0^{t+1}}$ for $t \geq 0$, so we need $|e_{T-1}| \geq \frac{1}{\varepsilon}$. Then $e_0 = \frac{1}{a_0}$ and:

$$e_t = \sum_{k=0}^{t-1} (-1)^k \frac{a_{k+1}}{a_0} e_{t-k-1},$$

for $t > 0$. To ensure $|e_t| \to \infty$ as $t \to \infty$, we need at least one of the roots of the infinite polynomial $p(z) = \sum_{k=0}^{\infty} (-z)^k \frac{a_k}{a_0}$ to lie strictly inside the unit circle. As $p(0) = 1$, it suffices that $0 > p(-1) = \sum_{k=0}^{\infty} \frac{a_k}{a_0}$. Since $a_0 > 0$, we just need $\sum_{k=0}^{\infty} a_k < 0$, i.e. for the the sum of the IRF of i_t to an anticipated monetary policy shock in the very far future to be negative.

To see whether this holds, we need to solve the original model without the ZLB but with a forcing process:

$$\pi_t = \kappa y_t + \beta\pi_{t+1},$$
$$y_t = y_{t+1} - \sigma^{-1}(i_t - \pi_{t+1} + \log\beta),$$
$$i_t = -\log\beta + \phi_\pi\pi_t + \phi_y y_t + \nu_t,$$

where unlike in the main text, we do not assume that $\nu_t = 0$ for $t \neq 1$. We guess that:

$$\pi_t = \sum_{j=0}^{\infty} F_j \nu_{t+j}.$$

So:

$$y_t = \frac{1}{\kappa}F_0\nu_t + \frac{1}{\kappa}\sum_{j=1}^{\infty}(F_j - \beta F_{j-1})\nu_{t+j},$$

$$i_t = -\log\beta + \left[1 + \left(\phi_\pi + \frac{\phi_y}{\kappa}\right)F_0\right]\nu_t + \sum_{j=1}^{\infty}\left[\phi_\pi F_j + \frac{\phi_y}{\kappa}(F_j - \beta F_{j-1})\right]\nu_{t+j},$$

$$\left[1 + \left(\phi_\pi + \frac{\phi_y}{\kappa} + \frac{\sigma}{\kappa}\right)F_0\right]\nu_t + \left[\phi_\pi F_1 + \frac{\phi_y}{\kappa}(F_1 - \beta F_0) - F_0 + \frac{\sigma}{\kappa}(F_1 - (1+\beta)F_0)\right]\nu_{t+1}$$
$$+ \sum_{j=2}^{\infty}\left[\phi_\pi F_j + \frac{\phi_y}{\kappa}(F_j - \beta F_{j-1}) - F_{j-1} + \frac{\sigma}{\kappa}(F_j - (1+\beta)F_{j-1} + \beta F_{j-2})\right]\nu_{t+j} = 0,$$

Matching terms gives:

$$F_0 = -\frac{\kappa}{\sigma + \kappa\phi_\pi + \phi_y},$$
$$F_1 = -\frac{\kappa\left(\kappa + \beta\phi_y + (1+\beta)\sigma\right)}{\left(\sigma + \kappa\phi_\pi + \phi_y\right)^2},$$
$$F_j = \frac{\left(\kappa + \beta\phi_y + (1+\beta)\sigma\right)F_{j-1} - \beta\sigma F_{j-2}}{\sigma + \kappa\phi_\pi + \phi_y},$$

for $j \geq 2$. From standard results on stationarity of $AR(2)$ processes, we have that $F_j \to 0$ as $j \to \infty$ if and only if the following three conditions are met:

$$\frac{\beta\sigma}{\sigma + \kappa\phi_\pi + \phi_y} < 1$$
$$\frac{\sigma + \kappa + \beta\phi_y}{\sigma + \kappa\phi_\pi + \phi_y} < 1$$
$$\frac{-\beta\sigma - \left(\kappa + \beta\phi_y + (1+\beta)\sigma\right)}{\sigma + \kappa\phi_\pi + \phi_y} < 1$$

The first always holds as $\beta < 1$ and other parameters are non-negative. The second follows from our assumption (for determinacy without the ZLB) that:

$$\kappa(\phi_\pi - 1) + (1-\beta)\phi_y > 0.$$

The third always holds as the numerator is negative and the denominator is positive. Thus, $F_j \to 0$ as $j \to \infty$. Moreover, this convergence is geometric.

From this solution for F_j we can infer a_j, since a_j gives the response of i_t to a shock anticipated in j periods. I.e.:

$$a_0 = 1 + \left(\phi_\pi + \frac{\phi_y}{\kappa}\right)F_0 = \frac{\sigma}{\sigma + \kappa\phi_\pi + \phi_y},$$
$$a_1 = -\frac{\phi_\pi\kappa(\kappa + (1+\beta)\sigma) + \phi_y(\kappa + \sigma)}{\left(\sigma + \kappa\phi_\pi + \phi_y\right)^2},$$

$$a_j = \phi_\pi F_j + \frac{\phi_y}{\kappa}\left(F_j - \beta F_{j-1}\right),$$

for $j \geq 1$. Now let $b_j = a_j + b_{j-1}$, with $b_0 = a_0$. So:

$$b_j = \phi_\pi F_j + \frac{\phi_y}{\kappa}\left(F_j - \beta F_{j-1}\right) + b_{j-1}.$$

$\lim_{j\to\infty} b_j$ exists and is finite as F_j converges geometrically. We just need to prove that $\lim_{j\to\infty} b_j < 0$. Since the calculations involved in this rapidly turn messy, we turn to Maple (worksheet available on request). Maple gives that:

$$\lim_{j\to\infty} b_j = -\frac{\kappa}{\kappa(\phi_\pi - 1) + (1-\beta)\phi_y} < 0$$

as we assumed $\kappa > 0$ and $\kappa(\phi_\pi - 1) + (1-\beta)\phi_y > 0$.

This establishes that as $T \to \infty$, $\frac{\left|\det M_{1:T-1,2:T}\right|}{a_0^T} = |e_T| \to \infty$. Hence, for any $\varepsilon > 0$, for sufficiently large T, $\frac{\left|\det M_{1:T-1,2:T}\right|}{a_0^T} \geq \frac{1}{\varepsilon}$, so our proposed $\pm\varepsilon$ change to the M matrix is sufficient to push its determinant negative.

Appendix H.4: Proof of Proposition 2

We first establish the following Lemma:

Lemma 4 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.

First, define $G := -C(B + CF)^{-1}$, and note that if L is the lag (right-shift) operator, the model from Problem 1 (Linear) 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 (Linear) 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 $\rho_{\mathcal{M}}$ for the spectral radius of some matrix $\mathcal{M}$, then, by this discussion and Assumption 2, $\rho_G < 1$.

Now consider the time reversed model:

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

subject to the terminal condition that $d_k \to 0$ as $k \to \infty$. 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 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. This completes the proof of Lemma 4.

Now 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:

$$s_t^* = [0_{n\times(t-1)} \quad s_1^*] = L^{t-1}s_1^*.$$

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 $t \in \mathbb{N}^+$, $M_{t,k} = \mathrm{O}(k^n \rho_G^k)$, 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. We need to provide conditions under which $\sup_{y\in[0,1]^{\mathbb{N}^+}} \inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y > 0$.[45]

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

$$\mathcal{C}_{\mathcal{M},\phi} := \sup_{k\in\mathbb{N}} \left\|(\mathcal{M}\phi^{-1})^k\right\|_2.$$

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

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

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

$$\rho_{\mathcal{M},\epsilon^*(\mathcal{M})} = 1.$$

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

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

[45] We might ideally have liked a lower bound on $\sup_{y\in\ell_1\cap[0,1]^{\mathbb{N}^+}} \inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y$ since by the Moore-Osgood theorem, this would imply a lower bound on $\lim_{T\to\infty} \max_{y\in[0,1]^T} \min_{t\in\{1,\dots,T\}} M_{t,1:T}y$ and thus imply that M was an S-matrix for all sufficiently large T. However, we have not managed to obtain a non-trivial lower bound on $\sup_{y\in\ell_1\cap[0,1]^{\mathbb{N}^+}} \inf_{t\in\mathbb{N}^+} M_{t,1:\infty}y$.

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

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

meaning $\mathcal{C}_{\mathcal{M},\phi} = \sup_{k=0,\dots,K}\left\|(\mathcal{M}\phi^{-1})^k\right\|_2$. The quantity $\rho_{\mathcal{M},\epsilon}$ (and hence $\epsilon^*(\mathcal{M})$) may be efficiently computed using the methods described by Wright and Trefethen (2001), and implemented in their EigTool toolkit[46]. Thus, given $\mathcal{M}$ and ϕ, $\mathcal{C}_{\mathcal{M},\phi}$ may be calculated in finitely many operations by iterating over $K \in \mathbb{N}$ until a K is found which satisfies:

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

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

$$\left\|\mathcal{M}^k\right\|_2 \le \mathcal{C}_{\mathcal{M},\phi}\phi^k.$$

Now, fix $\phi_F \in (\rho_F, 1)$ and $\phi_G \in (\rho_G, 1)$,[47] and define:

$$\mathcal{D}_{\phi_F,\phi_G} := \mathcal{C}_{F,\phi_F}\mathcal{C}_{G,\phi_F}\left\|(B+CF)^{-1}\right\|_2,$$

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

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

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

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

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

[47] In practice, we try a grid of values, as it is problem dependent whether high ϕ_F and low $\mathcal{C}_{F,\phi_F}$ is preferable to low ϕ_F and high $\mathcal{C}_{F,\phi_F}$.

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

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

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:

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

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.

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). Likewise, since $d_k \to 0$ as $k \to \infty$, as we have already proved that the first row of M converges to zero, by Lemma 4, it must be the case that $d_k = Hd_{k-1}$, for all $k > 0$, where $H = -(B + AH)^{-1}C$, and $\phi_H < 1$.

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:

$$\left|d_{-t,1}\right| \le \|d_{-t}\|_2 = \|F^t d_0\|_2 \le \|F^t\|_2 \|d_0\|_2 \le C_{F,\phi_F} \phi_F^t \|d_0\|_2,$$

and:

$$\left|d_{t,1}\right| \le \|d_t\|_2 = \|H^t d_0\|_2 \le \|H^t\|_2 \|d_0\|_2 \le C_{H,\phi_H} \phi_H^t \|d_0\|_2.$$

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 $\min\{t,k\} > T$, $\underline{M}^{(T)}_{t,k} = d_{k-t,1} - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G}$, then:

$$\varsigma \geq \max_{\substack{y \in [0,1]^T \\ y_\infty \in [0,1]}} \inf_{t \in \mathbb{N}^+} M_{t,1:\infty} \begin{bmatrix} y \\ y_\infty 1_{\infty \times 1} \end{bmatrix} \geq \max_{\substack{y \in [0,1]^T \\ y_\infty \in [0,1]}} \inf_{t \in \mathbb{N}^+} \underline{M}^{(T)}_{t,1:\infty} \begin{bmatrix} y \\ y_\infty 1_{\infty \times 1} \end{bmatrix}$$

$$= \max_{\substack{y \in [0,1]^T \\ y_\infty \in [0,1]}} \min \left\{ \begin{array}{c} \min_{t=1,\dots,T} \left[M_{t,1:T} y + \sum_{k=T+1}^{\infty} \left(d_{k-t,1} - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_\infty \right], \\ \inf_{t \in \mathbb{N}^+, t > T} \left[\sum_{k=1}^{T} \left(d_{k-t,1} - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_k + \sum_{k=T+1}^{\infty} \left(d_{k-t,1} - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_\infty \right] \end{array} \right\}$$

$$\geq \max_{\substack{y \in [0,1]^T \\ y_\infty \in [0,1]}} \min \left\{ \begin{array}{c} \min_{t=1,\dots,T} \left[M_{t,1:T} y + \left((I-H)^{-1} d_{T+1-t} \right)_1 y_\infty - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} y_\infty \right], \\ \min_{t=T+1,\dots,2T} \left[\begin{array}{c} \sum_{k=1}^{T} \left(d_{-(t-k),1} - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G} \right) y_k + \left((I-F)^{-1} (d_{-1} - d_{-(t-T)}) \right)_1 y_\infty \\ + \left((I-H)^{-1} d_0 \right)_1 y_\infty - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^{T+1}}{(1-\phi_F\phi_G)(1-\phi_G)} y_\infty \end{array} \right], \\ \inf_{t \in \mathbb{N}^+, t > 2T} \left[\begin{array}{c} \sum_{k=1}^{T} d_{-(t-k),1} y_k + \left((I-F)^{-1} (d_{-1} - d_{-(t-T)}) \right)_1 y_\infty \\ + \left((I-H)^{-1} d_0 \right)_1 y_\infty - \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^{2T+1} \phi_G}{(1-\phi_F\phi_G)(1-\phi_G)} \end{array} \right] \end{array} \right\}.$$

Now, for $t \geq T$:

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

so:

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

Thus $\varsigma \geq \underline{\varsigma}_T$, where:

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

The final minimand in this expression is less than (but converges to):

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

i.e. a weakly positive multiple of the doubly infinite sum of $d_{1,k}$ over all $k \in \mathbb{Z}$. If this expression is negative, then the optimum will have $y_\infty = 0$ giving (uninformatively) $\underline{\varsigma}_T \le 0$.

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 $\min\{t,k\} > T$, $\overline{M}^{(T)}_{t,k} = |d_{k-t,1}| + \mathcal{D}_{\phi_F,\phi_G} \frac{\phi_F^t \phi_G^k}{1-\phi_F\phi_G}$. Then:

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

Note that if $M_{1:T,1:T}$ is an S-matrix, $\overline{\varsigma}_T > 0$.

Appendix H.5: Proof of Proposition 3

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 (News) 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 (News), 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.

Appendix H.6: Proof of Theorem 3

First note that without loss of generality, we may assume that $L(\mathcal{f}) = 1$, since multiplying $\mathcal{f}_{k,s,e}$ by a constant does not change the solution. Additionally, we may assume that $\mathcal{f}_{k,s,e}$ is linear in its third argument for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$. To see this, consider the augmented system of equations:

$$u_t = \varepsilon_t,$$

$$v_t = x_{t-1},$$

$$w_t = \mathcal{f}_{k,s,u_{t-1}}(v_{t-1}, v_t, x_t),$$

$$0 = \mathbb{E}_t w_{t+1}.$$

This new system has the same solution(s) for x_t as the original one. It is also Lipschitz and linear in $t+1$ terms as required. Thus, we can assume that there exists a Lipschitz function $\tilde{\mathcal{f}}_{k,s,e}: \mathcal{X}^2 \to \mathbb{R}^n$ and a matrix $\mathcal{Q}_{k,s,e} \in \mathbb{R}^{n\times n}$ such that for all $x_0, x_1, x_2 \in \mathcal{X}$:

$$\mathcal{f}_{k,s,e}(x_0, x_1, x_2) = \tilde{\mathcal{f}}_{k,s,e}(x_0, x_1) + \mathcal{Q}_{k,s,e} x_2.$$

We also have that $\max\left\{L\left(\tilde{\mathcal{f}}_{k,s,e}\right), \left\|\mathcal{Q}_{k,s,e}\right\|_2\right\} \le L\left(\mathcal{f}_{k,s,e}\right) \le 1$, for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$.

Let $\mathcal{U} := \prod_{k\in K, s\in\mathbb{N}^+, e\in\mathcal{E}}(\mathcal{D}_{k,s} \to \mathbb{R}^n)$. This is the vector space of K-vectors of sequences of $\mathcal{E}$-vectors of functions. Note that $g^{(\sigma,\delta)} = \left(g_{k,s,e}^{(\sigma,\delta)}\right)_{k\in K, s\in\mathbb{N}^+, e\in\mathcal{E}}$ is a zero of the function $\Phi^{(\sigma,\delta)}: \mathcal{U} \to \mathcal{U}$ defined by:

$$\begin{aligned}
&\Phi^{(\sigma,\delta)}(g)_{k,s,e}(x) \\
&= (1-\delta)\Bigg[\tilde{\mathcal{f}}_{k,s,e}\left(x, g_{k,s,e}(x)\right) \\
&\qquad + \mathcal{Q}_{k,s,e}\left[(1-\sigma) g_{k,s+1,0}\left(g_{k,s,e}(x)\right) + \sigma \int_{\mathcal{E}} g_{k,s+1,\varepsilon}\left(g_{k,s,e}(x)\right) dp(\varepsilon)\right]\Bigg] \\
&+\delta \sum_{l\in K} \omega_{k,l,s,e}(x) \Bigg[\tilde{\mathcal{f}}_{k,s,e}\left(x, g_{k,s,e}(x)\right) \\
&\qquad + \mathcal{Q}_{k,s,e}\left[(1-\sigma) g_{l,1,0}\left(g_{k,s,e}(x)\right) + \sigma \int_{\mathcal{E}} g_{l,1,\varepsilon}\left(g_{k,s,e}(x)\right) dp(\varepsilon)\right]\Bigg] \\
&= \tilde{\mathcal{f}}_{k,s,e}\left(x, g_{k,s,e}(x)\right) + \mathcal{Q}_{k,s,e} g_{k,s+1,0}\left(g_{k,s,e}(x)\right) \\
&+\sigma \mathcal{Q}_{k,s,e} \int_{\mathcal{E}} \left(g_{k,s+1,\varepsilon} - g_{k,s+1,0}\right)\left(g_{k,s,e}(x)\right) dp(\varepsilon) \\
&+\delta(1-\sigma)\mathcal{Q}_{k,s,e} \sum_{l\in K} \omega_{k,l,s,e}(x)\left(g_{l,1,0} - g_{k,s+1,0}\right)\left(g_{k,s,e}(x)\right)
\end{aligned}$$

$$+\delta\sigma Q_{k,s,e}\sum_{l\in K}\omega_{k,l,s,e}(x)\int_{\mathcal{E}}\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x)\right)dp(\varepsilon)$$

for all $g\in\mathcal{U}$, $k\in K$, $s\in\mathbb{N}^+$, $e\in\mathcal{E}$ and $x\in\mathcal{D}_{k,s}$, where the equality follows from the fact that $\sum_{l\in K}\omega_{k,l,s,e}(x)=\int_{\mathcal{E}}1\,dp(\varepsilon)=1$.

We want to use the known perfect foresight solutions in order to prove the existence of the sunspot rational expectations solution for sufficiently small σ and δ. A natural approach to a problem like this is to use an implicit function theorem on $\Phi^{(\sigma,\delta)}$. There are two main difficulties. Firstly, we need the existence of a function, which requires an implicit function theorem for infinite dimensional spaces. Secondly, since $f_{k,s,e}$ (and hence $\Phi^{(\sigma,\delta)}$) is not everywhere differentiable, we need an implicit function theorem that does not rely on differentiability. We provide one with the required properties in Lemma 5 (An Implicit Function Theorem). This is an immediate special case of the result of Exercise 5H.6 of Dontchev & Rockafellar (2014). The notation we use in the statement of the Lemma follows that work:

Lemma 5 (An Implicit Function Theorem) Let P, X, Y and Z be metric spaces, with X complete and $Y\subseteq Z$ with the subspace metric.

Let $\bar{p}\in P$, $\bar{x}\in X$. Let $f: P\times X\to Z$ and $F: X\to Y$, with $f(\bar{p},\bar{x})=F(\bar{x})=0$. (*Note: F will act as an approximation to f.*)

Assume subtraction is defined on Z, with $F(x)-f(p,x)\in Y$ for all $p\in P$ and $x\in X$, and that $0\in Y$, where $y-0=y$ for all $y\in Y$.

Define $g: P\times X\to Y$ by $g(p,x)=F(x)-f(p,x)$.

Let $\kappa,\mu,\nu\in(0,\infty)$ with $\kappa\mu<1$.

Suppose:

1) There exist neighbourhoods $U\ni\bar{x}$ and $V\ni 0$ such that F is a bijection from $U\subseteq X$ to $V\subseteq Y$, and for all $y_1,y_2\in V$ with $y_1\neq y_2$, $\frac{d_X\left(F^{-1}(y_1),F^{-1}(y_2)\right)}{d_Y(y_1,y_2)}\leq\kappa$.
2) There exists $r_a>0$ such that if $p\in P$, $x_1,x_2\in X$, with $x_1\neq x_2$, $d_P(p,\bar{p})\leq r_a$, $d_X(x_1,\bar{x})\leq r_a$, $d_X(x_2,\bar{x})\leq r_a$, then $\frac{d_Y(g(p,x_1),g(p,x_2))}{d_X(x_1,x_2)}\leq\mu$.
3) There exists $r_b>0$ such that if $p_1,p_2\in P$, $x\in X$, with $p_1\neq p_2$, $d_P(p_1,\bar{p})\leq r_b$, $d_P(p_2,\bar{p})\leq r_b$, $d_X(x,\bar{x})\leq r_b$, then $\frac{d_Y(g(p_1,x),g(p_2,x))}{d_P(p_1,p_2)}\leq\nu$.

Then for all $\gamma > \frac{\kappa\nu}{1-\kappa\mu}$, there exists a neighbourhood $Q_\gamma \ni \bar{p}$ such that for all $p \in Q_\gamma \subseteq P$ there exists $x \in X$ with $f(p,x) = 0$ such that $d_X(\bar{x}, x) \le \gamma d_P(\bar{p}, p)$.

The proof of the underlying result is based on a contraction mapping theorem, and hence is constructive.

Our Φ will play the role of f in this Lemma (with an appropriately restricted domain), with g^* playing the role of $\bar{x}$.

Let $\mathcal{G} := \left\{g \in \mathcal{U} \middle| \|g - g^*\|_\infty < \infty\right\}$ and let $\mathcal{H} := \{h \in \mathcal{U} | \|h\|_\infty < \infty\}$, so $g^* \in \mathcal{G}$. $\mathcal{G}$ and $\mathcal{H}$ are complete metric spaces under the metric $d_{\mathcal{U}}(g,h) = \|g - h\|_\infty$ for $g, h \in \mathcal{U}$. Completeness for $\mathcal{H}$ follows from the fact that the space of bounded functions $\mathcal{D}_{k,s} \to \mathbb{R}^n$ is Banach for all $k \in K$, $s \in \mathbb{N}^+$, and from the fact that bounded norm, countable products of complete spaces are complete under the sup norm over the product. Completeness for $\mathcal{G}$ comes from the fact that $\mathcal{G} = \mathcal{H} + g^*$, with an identical metric.

Recall that our assumptions imply $A := A^{(0,0)} < \infty$ and $B = B^{(0,0)} < \infty$. So, for $g \in \mathcal{G}$, and $k, l \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s+1}$:

$$\begin{aligned}&\|g_{k,s+1,e}(x) - g_{k,s+1,0}(x)\|_2 \\ &\qquad = \left\|(g - g^*)_{k,s+1,e}(x) + g^*_{k,s+1,e}(x) - g^*_{k,s+1,0}(x) - (g - g^*)_{k,s+1,0}(x)\right\|_2,\end{aligned}$$

so:

$$\|g_{k,s+1,e}(x) - g_{k,s+1,0}(x)\|_2 \le A + 2\|g - g^*\|_\infty < \infty, \tag{24}$$

and likewise:

$$\|g_{l,1,e}(x) - g_{k,s+1,e}(x)\|_2 \le B + 2\|g - g^*\|_\infty < \infty. \tag{25}$$

Thus, to ensure $A^{(\sigma,\delta)}$ and $B^{(\sigma,\delta)}$ are finite, it suffices to ensure $g^{(\sigma,\delta)} \in \mathcal{G}$.

We now proceed with the definitions of the key spaces. The space X will be a closed (hence complete) subspace of a closed subspace W of $\mathcal{G}$, with the subspace metric. The spaces Y and Z will be closed subspaces of $\mathcal{H}$, again with the subspace metric. In particular, we fix:

$$c \in (0, C^{-1}),$$

then define:

$$\kappa := C(1 - cC)^{-1},$$

$$\chi := (1 + \kappa c)D,$$

and fix:

$$\hat{\lambda} \in (0, \min\{\lambda, \zeta\kappa^{-1}\}] \leq \lambda,$$
$$\mu \in (0, \kappa^{-1}),$$

and finally, define:

$$\hat{\zeta} := \kappa\hat{\lambda} \leq \zeta,$$
$$\overline{\sigma} := \min\left\{1, \frac{\hat{\lambda}}{2(A + 2\hat{\zeta})}, \frac{c}{4\chi^2}, \frac{\mu}{4(1+\chi)}\right\} > 0,$$
$$\bar{\delta} := \min\left\{1, \frac{\hat{\lambda}}{2(B + 2\hat{\zeta})}, \frac{c}{2[L_\Sigma(\omega)(B + 2\hat{\zeta}) + 2\chi^2]}, \frac{\mu}{4(1+\chi)}\right\} > 0,$$
$$W := \left\{g \in \mathcal{G} \middle| \|g - g^*\|_\infty \leq \hat{\zeta}, L(g) \leq \chi\right\},$$
$$Y := \{h \in \mathcal{H} | \|h\|_\infty \leq \hat{\lambda}, L(h) \leq c\},$$
$$Z := \{h \in \mathcal{H} | \|h\|_\infty \leq 2\hat{\lambda}, L(h) \leq 2c\}.$$

W, Y and Z are all closed as the set of functions with bounded Lipschitz norm is closed in the Banach space of bounded functions $\mathcal{D}_{k,s} \to \mathbb{R}^n$ for $k \in K$, $s \in \mathbb{N}^+$. Note that $g^* \in W$ as $L(g^*) \leq D \leq \chi$. We will define $X \subseteq W$ later.

We take the space $[0, \overline{\sigma}] \times [0, \bar{\delta}]$ to be the space P of parameters in the Lemma. The first coordinate gives σ and the second coordinate gives δ. We give this space the metric $d_P\big((\sigma_1, \delta_1), (\sigma_2, \delta_2)\big) = \max\{|\sigma_1 - \sigma_2|, |\delta_1 - \delta_2|\}$, for $\sigma_1, \sigma_2, \delta_1, \delta_2 \in [0,1]$. The point $(0,0)$ (i.e. $\sigma = 0$, $\delta = 0$) fulfils the role of $\bar{p}$ from the Lemma.

Let $\ell^\infty(\mathbb{N}^+, \mathbb{R}^n)$ be the Banach space of bounded $\mathbb{R}^n$ valued, $\mathbb{N}^+$ indexed, sequences, with the sup norm $\|\cdot\|_\infty$. Let $\mathcal{B}_{\hat{\lambda}} := \left\{\eta \in \ell^\infty(\mathbb{N}^+, \mathbb{R}^n) \middle| \|\eta\|_\infty \leq \hat{\lambda}\right\}$. This is a closed subset of $\ell^\infty(\mathbb{N}^+, \mathbb{R}^n)$, thus it is a complete metric space.

For $h \in Y$ (so $\|h\|_\infty \leq \hat{\lambda}$ and $L(h) \leq c$), and for $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$, we define the operator $\Theta^{(h)}_{k,s,e} \colon \mathcal{D}_{k,s} \to (\mathcal{B}_{\hat{\lambda}} \to \mathcal{B}_{\hat{\lambda}})$ by:

$$\Theta^{(h)}_{k,s,e}(x)(\eta) = \begin{pmatrix} h_{k,s,e}(x), \\ h_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right), \\ h_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right), \\ h_{k,s+3,0}\left(g^{*(R^2\eta)}_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right)\right), \\ \vdots \end{pmatrix},$$

for all $x \in \mathcal{D}_{k,s}$ and $\eta \in \mathcal{B}_{\hat{\lambda}}$. Note that for $x \in \mathcal{D}_{k,s}$ and $\eta_1, \eta_2 \in \mathcal{B}_{\hat{\lambda}}$:

$$\begin{aligned}
&\left\|\Theta^{(h)}_{k,s,e}(x)(\eta_1) - \Theta^{(h)}_{k,s,e}(x)(\eta_2)\right\|_\infty \\
&\leq L(h) \sup \left\{ \begin{array}{c}
\left\|g^{*(\eta_1)}_{k,s,e}(x) - g^{*(\eta_2)}_{k,s,e}(x)\right\|_2, \\
\left\|g^{*(R\eta_1)}_{k,s+1,0}\left(g^{*(\eta_1)}_{k,s,e}(x)\right) - g^{*(R\eta_2)}_{k,s+1,0}\left(g^{*(\eta_2)}_{k,s,e}(x)\right)\right\|_2, \\
\left\|g^{*(R^2\eta_1)}_{k,s+2,0}\left(g^{*(R\eta_1)}_{k,s+1,0}\left(g^{*(\eta_1)}_{k,s,e}(x)\right)\right) - g^{*(R^2\eta_2)}_{k,s+2,0}\left(g^{*(R\eta_2)}_{k,s+1,0}\left(g^{*(\eta_2)}_{k,s,e}(x)\right)\right)\right\|_2, \\
\vdots
\end{array} \right\} \\
&\leq L(h) C \|\eta_1 - \eta_2\|_\infty \leq cC \|\eta_1 - \eta_2\|_\infty,
\end{aligned}$$

by our assumption of $\mathcal{P}^{(\eta)}_{k,s,e}(x)$ being C-Lipschitz in η.

Also note that for $x \in \mathcal{D}_{k,s}$ and $\eta \in \mathcal{B}_{\hat{\lambda}}$:

$$\begin{aligned}
R\Theta^{(h)}_{k,s,e}(x)(\eta) &= \left(\begin{array}{c}
h_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right), \\
h_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right), \\
h_{k,s+3,0}\left(g^{*(R^2\eta)}_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right)\right), \\
\vdots
\end{array} \right) \\
&= \Theta^{(h)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)(R\eta).
\end{aligned}$$

Furthermore, note that for $h_1, h_2 \in Y$, and for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$, $x \in \mathcal{D}_{k,s}$ and $\eta \in \mathcal{B}_{\hat{\lambda}}$:

$$\begin{aligned}
&\left\|\Theta^{(h_1)}_{k,s,e}(x)(\eta) - \Theta^{(h_2)}_{k,s,e}(x)(\eta)\right\|_\infty \\
&= \sup \left\{ \begin{array}{c}
\left\|(h_1 - h_2)_{k,s,e}(x)\right\|_2, \\
\left\|(h_1 - h_2)_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right\|_2, \\
\left\|(h_1 - h_2)_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right)\right\|_2, \\
\left\|(h_1 - h_2)_{k,s+3,0}\left(g^{*(R^2\eta)}_{k,s+2,0}\left(g^{*(R\eta)}_{k,s+1,0}\left(g^{*(\eta)}_{k,s,e}(x)\right)\right)\right)\right\|_2, \\
\vdots
\end{array} \right\} \\
&\leq \|h_1 - h_2\|_\infty.
\end{aligned}$$

Additionally, for $h \in Y$, and for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$, $x_1, x_2 \in \mathcal{D}_{k,s}$ and $\eta \in \mathcal{B}_{\hat{\lambda}}$:

$$
\begin{aligned}
&\left\|\Theta_{k,s,e}^{(\hbar)}(x_1)(\eta)-\Theta_{k,s,e}^{(\hbar)}(x_2)(\eta)\right\|_\infty\\
&\le L(\hbar)\sup\left\{\begin{matrix}\left\|g_{k,s,e}^{*(\eta)}(x_1)-g_{k,s,e}^{*(\eta)}(x_2)\right\|_2,\\ \left\|g_{k,s+1,0}^{*(R\eta)}\left(g_{k,s,e}^{*(\eta)}(x_1)\right)-g_{k,s+1,0}^{*(R\eta)}\left(g_{k,s,e}^{*(\eta)}(x_2)\right)\right\|_2,\\ \left\|g_{k,s+2,0}^{*(R^2\eta)}\left(g_{k,s+1,0}^{*(R\eta)}\left(g_{k,s,e}^{*(\eta)}(x_1)\right)\right)-g_{k,s+2,0}^{*(R^2\eta)}\left(g_{k,s+1,0}^{*(R\eta)}\left(g_{k,s,e}^{*(\eta)}(x_2)\right)\right)\right\|_2,\\ \vdots\end{matrix}\right\}\\
&\le L(\hbar)D\|x_1-x_2\|_2\le cD\|x_1-x_2\|_2.
\end{aligned}
$$

Therefore, if $\hbar\in Y$, $\Theta_{k,s,e}^{(\hbar)}(x)\colon\mathcal{B}_{\hat\lambda}\to\mathcal{B}_{\hat\lambda}$ is a contraction mapping for all $k\in K$, $s\in\mathbb{N}^+$, $e\in\mathcal{E}$ and $x\in\mathcal{D}_{k,s}$. Thus as $\mathcal{B}_{\hat\lambda}$ is complete, by the Banach fixed point theorem, $\Theta_{k,s,e}^{(\hbar)}(x)\colon\mathcal{B}_{\hat\lambda}\to\mathcal{B}_{\hat\lambda}$ has a unique fixed point, which we denote $\eta_{k,s,e}^{(\hbar)}(x)\in\mathcal{B}_{\hat\lambda}$. Clearly $\eta_{k,s,e}^{(0)}(x)=0$. Moreover, by the parametric version of the Banach fixed point theorem given in Theorem 1A.4 of Dontchev & Rockafellar (2014), in fact $\eta_{k,s,e}^{(\hbar)}(x)$ is Lipschitz continuous in $\hbar$ (with k,s,e,x fixed), with Lipschitz modulus $(1-cC)^{-1}$, and $\eta_{k,s,e}^{(\hbar)}(x)$ is Lipschitz continuous in x (with $\hbar,k,s,e$ fixed), with Lipschitz modulus $cD(1-cC)^{-1}$.

Furthermore, by the definition of $g^{*(\eta)}$, we have that:

$$
\hbar_{k,s,e}\left(x,g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x),g_{k,s+1,0}^{*\left(R\eta_{k,s,e}^{(\hbar)}(x)\right)}\left(g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x)\right)\right)=\eta_{k,s,e}^{(\hbar)}(x)_1,
$$

so as $\eta_{k,s,e}^{(\hbar)}(x)=\Theta_{k,s,e}^{(\hbar)}(x)\left(\eta_{k,s,e}^{(\hbar)}(x)\right)$ and:

$$
R\eta_{k,s,e}^{(\hbar)}(x)=R\Theta_{k,s,e}^{(\hbar)}(x)\left(\eta_{k,s,e}^{(\hbar)}(x)\right)=\Theta_{k,s+1,0}^{(\hbar)}\left(g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x)\right)\left(R\eta_{k,s,e}^{(\hbar)}(x)\right),
$$

implying $\eta_{k,s+1,0}^{(\hbar)}\left(g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x)\right)=R\eta_{k,s,e}^{(\hbar)}(x)$:

$$
\hbar_{k,s,e}\left(x,g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x),g_{k,s+1,0}^{*\left(\eta_{k,s+1,0}^{(\hbar)}\left(g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x)\right)\right)}\left(g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x)\right)\right)=\hbar_{k,s,e}(x).
$$

We now define $\hat g^{*(\hbar)}$ by:

$$
\hat g_{k,s,e}^{*(\hbar)}(x)=g_{k,s,e}^{*\left(\eta_{k,s,e}^{(\hbar)}(x)\right)}(x),
$$

for all $k\in K$, $s\in\mathbb{N}^+$, $e\in\mathcal{E}$ and $x\in\mathcal{D}_{k,s}$. With this definition, the previous equality can be restated as:

$$f_{k,s,e}\left(x, \hat{g}^{*(h)}_{k,s,e}(x), \hat{g}^{*(h)}_{k,s+1,0}\left(\hat{g}^{*(h)}_{k,s,e}(x)\right)\right) = h_{k,s,e}(x). \tag{26}$$

Additionally, note that for $h_1, h_2 \in Y$, and for all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$:

$$\begin{aligned}\left\|\hat{g}^{*(h_1)}_{k,s,e}(x) - \hat{g}^{*(h_2)}_{k,s,e}(x)\right\|_2 &= \left\|g^{*\left(\eta^{(h_1)}_{k,s,e}(x)\right)}_{k,s,e}(x) - g^{*\left(\eta^{(h_2)}_{k,s,e}(x)\right)}_{k,s,e}(x)\right\|_2 \\ &\le C\left\|\eta^{(h_1)}_{k,s,e}(x) - \eta^{(h_2)}_{k,s,e}(x)\right\|_\infty \\ &\le C(1 - cC)^{-1}\|h_1 - h_2\|_\infty\end{aligned}$$

where the first inequality comes from our assumption on $g^{*(\eta)}$, and the latter comes from our previous result that $\eta^{(h)}_{k,s,e}(x)$ is $(1 - cC)^{-1}$-Lipschitz in h. Thus as $\kappa = C(1 - cC)^{-1}$:

$$\left\|\hat{g}^{*(h_1)} - \hat{g}^{*(h_2)}\right\|_\infty \le \kappa\|h_1 - h_2\|_\infty,$$

i.e., $\hat{g}^{*(h)}$ is κ-Lipschitz in h. Since $\hat{g}^{*(0)} = g^*$ (as $\eta^{(0)}_{k,s,e}(x) = 0$), setting $h_2 = 0$ in this relationship implies that for $h_1 \in Y$,

$$\left\|\hat{g}^{*(h_1)} - g^*\right\|_\infty \le \kappa\|h_1\|_\infty \le \kappa\hat{\lambda} = \hat{\zeta}.$$

Furthermore, since for $h \in Y$, $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x_1, x_2 \in \mathcal{D}_{k,s}$:

$$\begin{aligned}&\left\|\hat{g}^{*(h)}_{k,s,e}(x_1) - \hat{g}^{*(h)}_{k,s,e}(x_2)\right\|_2 \\ &= \left\|g^{*\left(\eta^{(h)}_{k,s,e}(x_1)\right)}_{k,s,e}(x_1) - g^{*\left(\eta^{(h)}_{k,s,e}(x_1)\right)}_{k,s,e}(x_2) + g^{*\left(\eta^{(h)}_{k,s,e}(x_1)\right)}_{k,s,e}(x_2) - g^{*\left(\eta^{(h)}_{k,s,e}(x_2)\right)}_{k,s,e}(x_2)\right\|_2 \\ &\le \left\|g^{*\left(\eta^{(h)}_{k,s,e}(x_1)\right)}_{k,s,e}(x_1) - g^{*\left(\eta^{(h)}_{k,s,e}(x_1)\right)}_{k,s,e}(x_2)\right\|_2 + \left\|g^{*\left(\eta^{(h)}_{k,s,e}(x_1)\right)}_{k,s,e}(x_2) - g^{*\left(\eta^{(h)}_{k,s,e}(x_2)\right)}_{k,s,e}(x_2)\right\|_2 \\ &\le D\|x_1 - x_2\|_2 + C\left\|\eta^{(h)}_{k,s,e}(x_1) - \eta^{(h)}_{k,s,e}(x_2)\right\|_2 \le D\|x_1 - x_2\|_2 + cCD(1 - cC)^{-1}\|x_1 - x_2\|_2,\end{aligned}$$

as $\eta^{(h)}_{k,s,e}(x)$ is $cD(1 - cC)^{-1}$-Lipschitz in x. Thus, $L(\hat{g}^{*(h)}) \le D + cCD(1 - cC)^{-1} = \chi$. This establishes that for all $h \in Y$, $h \in W$. We set:

$$X := \hat{g}^{*(Y)} := \{\hat{g}^{*(h)} | h \in Y\}.$$

Hence, we have that $X \subseteq W$. We still need to prove that X is closed as required for completeness.

Next, we define the operator $\Psi: W \to \mathcal{U}$ by:

$$\begin{aligned}\Psi(g)_{k,s,e}(x) &= f_{k,s,e}\left(x, g_{k,s,e}(x), g_{k,s+1,0}\left(g_{k,s,e}(x)\right)\right) \\ &= \tilde{f}_{k,s,e}\left(x, g_{k,s,e}(x)\right) + Q_{k,s,e}g_{k,s+1,0}\left(g_{k,s,e}(x)\right),\end{aligned}$$

for all $g \in W$, $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$. We write $a|_S$ for the restriction of (the domain of) some function a to some set S. $\Psi|_X$ will play the role of F in the Lemma.

Now, for $g_1, g_2 \in W$, $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$:

$$\begin{aligned}
&\left\|\Psi\left(g_1\right)_{k,s,e}(x) - \Psi\left(g_2\right)_{k,s,e}(x)\right\|_2 \\
&\leq L(\tilde{f}_{k,s,e})\left\|g_{1,k,s,e}(x) - g_{2,k,s,e}(x)\right\|_2 \\
&+\|Q_{k,s,e}\|_2\left\|g_{1,k,s+1,0}\left(g_{1,k,s,e}(x)\right) - g_{2,k,s+1,0}\left(g_{1,k,s,e}(x)\right) + g_{2,k,s+1,0}\left(g_{1,k,s,e}(x)\right)\right. \\
&\qquad\qquad \left. - g_{2,k,s+1,0}\left(g_{2,k,s,e}(x)\right)\right\|_2 \\
&\leq L(\tilde{f}_{k,s,e})\left\|g_1 - g_2\right\|_\infty + \|Q_{k,s,e}\|_2\left[\left\|g_1 - g_2\right\|_\infty + L\left(g_2\right)\left\|g_1 - g_2\right\|_\infty\right] \\
&\leq [2+\chi]\left\|g_1 - g_2\right\|_\infty,
\end{aligned}$$

where the first inequality follows from the triangle inequality, the definition of the matrix norm, and the fact that $\tilde{f}_{k,s,e}$ is Lipschitz, the second comes from the triangle inequality and the fact that g_2 is Lipschitz and the third comes from the fact that $\max\left\{L\left(\tilde{f}_{k,s,e}\right), \left\|Q_{k,s,e}\right\|_2\right\} \leq 1$. Thus Ψ is $[2+\chi]$-Lipschitz. With $g_2 = g^* \in W$, this implies that for all $g_1 \in W$, $\left\|\Psi\left(g_1\right)\right\|_\infty \leq [2+\chi]\hat{\zeta}$, so $\Psi\left(g_1\right) \in \mathcal{H}$. Hence as $X = \hat{g}^{*(Y)} \subseteq W$, $\Psi|_X: X \to \mathcal{H}$ is continuous. Thus as Y is closed in $\mathcal{H}$, by continuity $(\Psi|_X)^{-1}(Y) \subseteq W$ is closed in W and thus closed in $\mathcal{G}$.

Now note that for all $h \in Y$, and all $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$:

$$\Psi\left(\hat{g}^{*(h)}\right)_{k,s,e}(x) = f_{k,s,e}\left(x, \hat{g}^{*(h)}_{k,s,e}(x), \hat{g}^{*(h)}_{k,s+1,0}\left(\hat{g}^{*(h)}_{k,s,e}(x)\right)\right) = h_{k,s,e}(x),$$

by our result in equation (26). I.e., $\Psi\left(\hat{g}^{*(h)}\right) = h$. Thus if $g \in X$, then by the definition of X, there exists $h \in Y$ such that $g = \hat{g}^{*(h)}$ and $\Psi\left(g\right) = \Psi\left(\hat{g}^{*(h)}\right) = h \in Y$. Hence, $\Phi(X) = Y$. Similarly, if for $g_1, g_2 \in X$, $\Psi(g_1) = \Psi(g_2)$, then $g_1 = g_2 = \hat{g}^{*\left(\Psi(g_1)\right)}$. Thus $\Psi|_X$ is a bijection from X to Y. Thus, as we already established that $(\Psi|_X)^{-1}(Y)$ is closed in $\mathcal{G}$, X is closed in $\mathcal{G}$ as required for completeness. Moreover, $\Psi|_X: X \to Y$ is $[2+\chi]$-Lipschitz and $(\Psi|_X)^{-1}: Y \to X$ is κ-Lipschitz, with $(\Psi|_X)^{-1}(h) = \hat{g}^{*(h)}$ for all $h \in Y$.

Then, for the first condition of the Lemma, we claim we can take $U := X$, $V := Y$ and $\kappa = C(1-cC)^{-1}$ as already defined. $F = \Psi|_X$ is a κ-Lipschitz bijection from U to V. The only thing left to check is that U is a neighbourhood of g^* and V is a neighbourhood of 0. Let $Y_0 := \{h \in Y \mid \|h\|_\infty < \hat{\lambda}\} \subseteq Y = V$. Y_0 is the intersection of a set open in $\mathcal{H}$ with Y, thus Y_0 is open in Y. Let $X_0 := (\Psi|_X)^{-1}(Y_0) \subseteq X = U$. X_0 is open in X as $\Psi|_X: X \to Y$ is continuous and Y_0 is open in Y. Since $0 \in Y_0 \subseteq V$, $g^* \in X_0 \subseteq U$, as required. This establishes that U is a neighbourhood of g^* and V is a neighbourhood of 0, completing the proof that the Lemma's first condition holds.

Following the Lemma, we now define the operator $\Gamma^{(\sigma,\delta)}: X \to \mathcal{U}$ for $\sigma < \overline{\sigma}$ and $\delta < \bar{\delta}$ by:

$$\Gamma^{(\sigma,\delta)}(g) = \Phi^{(\sigma,\delta)}(g) - \Psi(g),$$

for all $g \in X$. Γ will play the role of $-g$ in the Lemma. (The sign makes no difference in the below.) For $k \in K$, $s \in \mathbb{N}^+$, $e \in \mathcal{E}$ and $x \in \mathcal{D}_{k,s}$:

$$\begin{aligned}\Gamma^{(\sigma,\delta)}(g)_{k,s,e}(x) = {} & \sigma Q_{k,s,e} \int_{\mathcal{E}} (g_{k,s+1,\varepsilon} - g_{k,s+1,0})\big(g_{k,s,e}(x)\big)\, dp(\varepsilon) \\ & +\delta(1-\sigma) Q_{k,s,e} \sum_{l \in K} \omega_{k,l,s,e}(x) (g_{l,1,0} - g_{k,s+1,0})\big(g_{k,s,e}(x)\big) \\ & +\delta\sigma Q_{k,s,e} \sum_{l \in K} \omega_{k,l,s,e}(x) \int_{\mathcal{E}} (g_{l,1,\varepsilon} - g_{k,s+1,\varepsilon})\big(g_{k,s,e}(x)\big)\, dp(\varepsilon).\end{aligned}$$

We want to show that $\Gamma^{(\sigma,\delta)}(X) \subseteq Y$. Let $g \in X$, $k \in K$, $s \in \mathbb{N}^+$ and $e \in \mathcal{E}$. We need to prove that $\left\|\Gamma^{(\sigma,\delta)}(g)_{k,s,e}\right\|_\infty \le \hat{\lambda}$ and $L\left(\Gamma^{(\sigma,\delta)}(g)_{k,s,e}\right) \le c$.

First let $x \in \mathcal{D}_{k,s}$. Then:

$$\begin{aligned}& \left\|\Gamma^{(\sigma,\delta)}(g)_{k,s,e}(x)\right\|_2 \\ & \le \sigma \|Q_{k,s,e}\|_2 \int_{\mathcal{E}} \left\|(g_{k,s+1,\varepsilon} - g_{k,s+1,0})\big(g_{k,s,e}(x)\big)\right\|_2 dp(\varepsilon) \\ & +\delta(1-\sigma)\|Q_{k,s,e}\|_2 \sum_{l \in K} \omega_{k,l,s,e}(x) \left\|(g_{l,1,0} - g_{k,s+1,0})\big(g_{k,s,e}(x)\big)\right\|_2 \\ & +\delta\sigma \|Q_{k,s,e}\|_2 \sum_{l \in K} \omega_{k,l,s,e}(x) \int_{\mathcal{E}} \left\|(g_{l,1,\varepsilon} - g_{k,s+1,\varepsilon})\big(g_{k,s,e}(x)\big)\right\|_2 dp(\varepsilon) \\ & \le \sigma \int_{\mathcal{E}} \left[A + 2\|g - g^*\|_\infty\right] dp(\varepsilon) + \delta(1-\sigma) \sum_{l \in K} \omega_{k,l,s,e}(x) \left[B + 2\|g - g^*\|_\infty\right] \\ & \qquad + \delta\sigma \sum_{l \in K} \omega_{k,l,s,e}(x) \int_{\mathcal{E}} \left[B + 2\|g - g^*\|_\infty\right] dp(\varepsilon) \\ & \le \sigma\big(A + 2\hat{\zeta}\big) + \delta\big(B + 2\hat{\zeta} + B\big) \le \frac{\hat{\lambda}}{2} + \frac{\hat{\lambda}}{2} = \hat{\lambda}.\end{aligned}$$

The first inequality comes from the triangle inequality. The second comes from the results in equations (24) and (25) and the fact that $\|Q_{k,s,e}\|_2 \le 1$. The third comes from the fact that $\sum_{l \in K} \omega_{k,l,s,e}(x) = \int_{\mathcal{E}} 1\, dp(\varepsilon) = 1$ and from the definition of $W \supseteq X$. The fourth comes from the definition of $\overline{\sigma}$ and $\bar{\delta}$.

Now let $x_1, x_2 \in \mathcal{D}_{k,s}$. Then:

$$\begin{aligned}& \left\|\Gamma^{(\sigma,\delta)}(g)_{k,s,e}(x_1) - \Gamma^{(\sigma,\delta)}(g)_{k,s,e}(x_2)\right\|_2 \\ & \le \sigma \|Q_{k,s,e}\|_2 \int_{\mathcal{E}} \Big[\left\|g_{k,s+1,\varepsilon}\big(g_{k,s,e}(x_1)\big) - g_{k,s+1,\varepsilon}\big(g_{k,s,e}(x_2)\big)\right\|_2 \\ & \qquad + \left\|g_{k,s+1,0}\big(g_{k,s,e}(x_1)\big) - g_{k,s+1,0}\big(g_{k,s,e}(x_2)\big)\right\|_2\Big]\, \mathrm{d}p(\varepsilon)\end{aligned}$$

$$
\begin{aligned}
&+\delta(1-\sigma)\left\|Q_{k,s,e}\right\|_2 \Bigg\| \sum_{l\in K} \omega_{k,l,s,e}(x_1)\left(g_{l,1,0}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x_1)\right) \\
&\qquad - \sum_{l\in K} \omega_{k,l,s,e}(x_1)\left(g_{l,1,0}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x_2)\right) \\
&\qquad + \sum_{l\in K} \omega_{k,l,s,e}(x_1)\left(g_{l,1,0}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x_2)\right) \\
&\qquad - \sum_{l\in K} \omega_{k,l,s,e}(x_2)\left(g_{l,1,0}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x_2)\right)\Bigg\|_2 \\
&+\delta\sigma\left\|Q_{k,s,e}\right\|_2 \Bigg\| \sum_{l\in K} \omega_{k,l,s,e}(x_1) \int_{\mathcal{E}}\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x_1)\right) dp(\varepsilon) \\
&\qquad - \sum_{l\in K} \omega_{k,l,s,e}(x_1) \int_{\mathcal{E}}\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x_2)\right) dp(\varepsilon) \\
&\qquad + \sum_{l\in K} \omega_{k,l,s,e}(x_1) \int_{\mathcal{E}}\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x_2)\right) dp(\varepsilon) \\
&\qquad - \sum_{l\in K} \omega_{k,l,s,e}(x_2) \int_{\mathcal{E}}\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x_2)\right) dp(\varepsilon)\Bigg\|_2 \\
&\leq \sigma \int_{\mathcal{E}} 2L(g)\left\|g_{k,s,e}(x_1)-g_{k,s,e}(x_2)\right\|_2 dp(\varepsilon) \\
&+\delta(1-\sigma)\sum_{l\in K} \omega_{k,l,s,e}(x_1)\Big[\left\|g_{l,1,0}\left(g_{k,s,e}(x_1)\right)-g_{l,1,0}\left(g_{k,s,e}(x_2)\right)\right\|_2 \\
&\qquad + \left\|g_{k,s+1,0}\left(g_{k,s,e}(x_1)\right)-g_{k,s+1,0}\left(g_{k,s,e}(x_2)\right)\right\|_2\Big] \\
&+\delta(1-\sigma)\sum_{l\in K}\left|\omega_{k,l,s,e}(x_1)-\omega_{k,l,s,e}(x_2)\right|\left\|\left(g_{l,1,0}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x_2)\right)\right\|_2 \\
&+\delta\sigma\sum_{l\in K} \omega_{k,l,s,e}(x_1)\int_{\mathcal{E}}\Big[\left\|g_{l,1,\varepsilon}\left(g_{k,s,e}(x_1)\right)-g_{l,1,\varepsilon}\left(g_{k,s,e}(x_2)\right)\right\|_2 \\
&\qquad + \left\|g_{k,s+1,\varepsilon}\left(g_{k,s,e}(x_1)\right)-g_{k,s+1,\varepsilon}\left(g_{k,s,e}(x_2)\right)\right\|_2\Big] dp(\varepsilon) \\
&+\delta\sigma\sum_{l\in K}\left|\omega_{k,l,s,e}(x_1)-\omega_{k,l,s,e}(x_2)\right|\int_{\mathcal{E}}\left\|\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x_2)\right)\right\|_2 dp(\varepsilon) \\
&\leq \sigma\int_{\mathcal{E}} 2L(g)^2\|x_1-x_2\|_2\, dp(\varepsilon) \\
&+\delta(1-\sigma)\sum_{l\in K}\omega_{k,l,s,e}(x_1)2L(g)^2\|x_1-x_2\|_2 \\
&+\delta(1-\sigma)\sum_{l\in K}L(\omega_{k,l,s,e})\|x_1-x_2\|_2\left[B+2\|g-g^*\|_\infty\right] \\
&+\delta\sigma\sum_{l\in K}\omega_{k,l,s,e}(x_1)\int_{\mathcal{E}}2L(g)^2\|x_1-x_2\|_2\, dp(\varepsilon) \\
&+\delta\sigma\sum_{l\in K}L(\omega_{k,l,s,e})\|x_1-x_2\|_2\int_{\mathcal{E}}\left[B+2\|g-g^*\|_\infty\right] dp(\varepsilon) \\
&\leq \sigma(2\chi^2)\|x_1-x_2\|_2+\delta\left[L_\Sigma(\omega)\left(B+2\hat{\zeta}\right)+2\chi^2\right]\|x_1-x_2\|_2 \\
&\leq \frac{c}{2}\|x_1-x_2\|_2+\frac{c}{2}\|x_1-x_2\|_2\leq c\|x_1-x_2\|_2.
\end{aligned}
$$

The first inequality comes from the triangle inequality. The second comes from the triangle inequality, g being Lipschitz, and the fact that $\|Q_{k,s,e}\|_2 \leq 1$. The third comes from g and $\omega_{k,l,s,e}$ being Lipschitz, and from the result in equations (25). The fourth comes from the fact that

$\sum_{l\in K}\omega_{k,l,s,e}(x_1)=\int_{\mathcal{E}} 1\, dp(\varepsilon)=1$, and from the definitions of $L_\Sigma(\omega)$ and $W\supseteq X$. The fifth comes from the definitions of $\overline{\sigma}$ and $\bar{\delta}$. This completes the proof that $\Gamma^{(\sigma,\delta)}(X)\subseteq Y$. Additionally, note that for $g\in X$, $\Phi^{(\sigma,\delta)}(g)=\Psi(g)+\Gamma^{(\sigma,\delta)}(g)$. As $\Psi(g)\in Y$ and $\Gamma^{(\sigma,\delta)}(g)\in Y$, so by the definition of Z, $\Phi^{(\sigma,\delta)}(g)\in Z$, as required.

For the Lemma's second condition, let $g_1,g_2\in X$, $k\in K$, $s\in\mathbb{N}^+$, $x\in\mathcal{D}_{k,s}$ and $e\in\mathcal{E}$. Then:

$$\begin{aligned}
&\left\|\Gamma^{(\sigma,\delta)}(g_1)_{k,s,e}(x)-\Gamma^{(\sigma,\delta)}(g_2)_{k,s,e}(x)\right\|_2\\
&\le\sigma\int_{\mathcal{E}}\left\|(g_{1,k,s+1,\varepsilon}-g_{1,k,s+1,0})\left(g_{1,k,s,e}(x)\right)-(g_{1,k,s+1,\varepsilon}-g_{1,k,s+1,0})\left(g_{2,k,s,e}(x)\right)\right\|_2 dp(\varepsilon)\\
&+\sigma\int_{\mathcal{E}}\left\|(g_{1,k,s+1,\varepsilon}-g_{1,k,s+1,0})\left(g_{2,k,s,e}(x)\right)-(g_{2,k,s+1,\varepsilon}-g_{2,k,s+1,0})\left(g_{2,k,s,e}(x)\right)\right\|_2 dp(\varepsilon)\\
&+\delta(1-\sigma)\sum_{l\in K}\omega_{k,l,s,e}(x)\left\|(g_{1,l,1,0}-g_{1,k,s+1,0})\left(g_{1,k,s,e}(x)\right)\right.\\
&\qquad\qquad\left.-(g_{1,l,1,0}-g_{1,k,s+1,0})\left(g_{2,k,s,e}(x)\right)\right\|_2\\
&+\delta(1-\sigma)\sum_{l\in K}\omega_{k,l,s,e}(x)\left\|(g_{1,l,1,0}-g_{1,k,s+1,0})\left(g_{2,k,s,e}(x)\right)\right.\\
&\qquad\qquad\left.-(g_{2,l,1,0}-g_{2,k,s+1,0})\left(g_{2,k,s,e}(x)\right)\right\|_2\\
&+\delta\sigma\sum_{l\in K}\omega_{k,l,s,e}(x)\int_{\mathcal{E}}\left\|(g_{1,l,1,\varepsilon}-g_{1,k,s+1,\varepsilon})\left(g_{1,k,s,e}(x)\right)\right.\\
&\qquad\qquad\left.-(g_{1,l,1,\varepsilon}-g_{1,k,s+1,\varepsilon})\left(g_{2,k,s,e}(x)\right)\right\|_2 dp(\varepsilon)\\
&+\delta\sigma\sum_{l\in K}\omega_{k,l,s,e}(x)\int_{\mathcal{E}}\left\|(g_{1,l,1,\varepsilon}-g_{1,k,s+1,\varepsilon})\left(g_{2,k,s,e}(x)\right)\right.\\
&\qquad\qquad\left.-(g_{2,l,1,\varepsilon}-g_{2,k,s+1,\varepsilon})\left(g_{2,k,s,e}(x)\right)\right\|_2 dp(\varepsilon)\\
&\le\sigma\int_{\mathcal{E}}2L(g_1)\|g_1-g_2\|_\infty dp(\varepsilon)+\sigma\int_{\mathcal{E}}2\|g_1-g_2\|_\infty dp(\varepsilon)\\
&+\delta(1-\sigma)\sum_{l\in K}\omega_{k,l,s,e}(x)2L(g_1)\|g_1-g_2\|_\infty+\delta(1-\sigma)\sum_{l\in K}\omega_{k,l,s,e}(x)2\|g_1-g_2\|_\infty\\
&+\delta\sigma\sum_{l\in K}\omega_{k,l,s,e}(x)\int_{\mathcal{E}}2L(g_1)\|g_1-g_2\|_\infty dp(\varepsilon)+\delta\sigma\sum_{l\in K}\omega_{k,l,s,e}(x)\int_{\mathcal{E}}2\|g_1-g_2\|_\infty dp(\varepsilon)\\
&\le 2(\sigma+\delta)(1+\chi)\|g_1-g_2\|_\infty\le\mu\|g_1-g_2\|_\infty.
\end{aligned}$$

The first inequality comes from the triangle inequality and the fact that $\|Q_{k,s,e}\|_2\le 1$. The second comes from the triangle inequality and g_1 being Lipschitz. The third comes from the fact that $\sum_{l\in K}\omega_{k,l,s,e}(x_1)=\int_{\mathcal{E}} 1\, dp(\varepsilon)=1$, and from the definition of $W\supseteq X$. The fourth comes from the definitions of $\overline{\sigma}$ and $\bar{\delta}$. This establishes the Lemma's second condition for any arbitrarily large $r_a>0$.

For the Lemma's third and final condition, let $g\in X$, $k\in K$, $s\in\mathbb{N}^+$, $x\in\mathcal{D}_{k,s}$ and $e\in\mathcal{E}$. Then for $\sigma_1<\overline{\sigma}$, $\sigma_2<\overline{\sigma}$, $\delta_1<\bar{\delta}$ and $\delta_2<\bar{\delta}$:

$$
\begin{aligned}
&\left\|\Gamma^{(\sigma_1,\delta_1)}(g)_{k,s,e}(x)-\Gamma^{(\sigma_2,\delta_2)}(g)_{k,s,e}(x)\right\|_2\\
&\le|\sigma_1-\sigma_2|\int_{\mathcal{E}}\left\|\left(g_{k,s+1,\varepsilon}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x)\right)\right\|_2 dp(\varepsilon)\\
&+[|\delta_1(1-\sigma_1)-\delta_1(1-\sigma_2)|\\
&\qquad+|\delta_1(1-\sigma_2)-\delta_2(1-\sigma_2)|]\sum_{l\in K}\omega_{k,l,s,e}(x)\left\|\left(g_{l,1,0}-g_{k,s+1,0}\right)\left(g_{k,s,e}(x)\right)\right\|_2\\
&+[|\delta_1\sigma_1-\delta_1\sigma_2|+|\delta_1\sigma_2-\delta_2\sigma_2|]\sum_{l\in K}\omega_{k,l,s,e}(x)\int_{\mathcal{E}}\left\|\left(g_{l,1,\varepsilon}-g_{k,s+1,\varepsilon}\right)\left(g_{k,s,e}(x)\right)\right\|_2 dp(\varepsilon)\\
&\le|\sigma_1-\sigma_2|\left[A+2\|g-g^*\|_\infty\right]+[2\delta_1|\sigma_1-\sigma_2|+|\delta_1-\delta_2|]\left[B+2\|g-g^*\|_\infty\right]\\
&\le\left[A+2\hat{\zeta}+\left(1+2\bar{\delta}\right)\left(B+2\hat{\zeta}\right)\right]\max\{|\sigma_1-\sigma_2|,|\delta_1-\delta_2|\},
\end{aligned}
$$

where the first inequality follows from the triangle inequality, the second comes from the results in equations (24) and (25), and the fact that $\sum_{l\in K}\omega_{k,l,s,e}(x_1)=\int_{\mathcal{E}}1\,dp(\varepsilon)=1$, and the third comes from the definition of $W\supseteq X$. Thus, we can take:

$$\nu:=A+2\hat{\zeta}+\left(1+2\bar{\delta}\right)\left(B+2\hat{\zeta}\right),$$

and then the Lemma's third condition will be satisfied for arbitrarily large $r_b>0$.

Thus, from the conclusion of the lemma, for all $\gamma>\frac{\kappa\nu}{1-\kappa\mu}$, there exists:

$$\xi\in\left(0,\min\{\overline{\sigma},\bar{\delta}\}\right],$$

such that for all $\sigma<\xi$ and $\delta<\xi$, there exists $g^{(\sigma,\delta)}\in X$ with $\Phi^{(\sigma,\delta)}\left(g^{(\sigma,\delta)}\right)=0$ (i.e., solving the model) such that $\left\|g^{(\sigma,\delta)}-g^*\right\|_\infty\le\gamma\max\{|\sigma|,|\delta|\}$.

Appendix H.7: Proof of Corollary 5

For a given initial state $x_0\in\mathbb{R}^n$, and initial shock $\varepsilon_1\in\mathcal{E}$, we write $i_t(x_0,\varepsilon_1)$ and $x_t(x_0,\varepsilon_1)$ for the values of i_t and x_t, respectively, under the perfect foresight solution given those initial conditions, with $\varepsilon_t=0$ for $t>1$. The existence of such a solution is guaranteed by our assumption above.

Now let:

$$T(x_0,\varepsilon_1):=\min\{T\in\mathbb{N}|\forall t>T,i_t(x_0,\varepsilon_1)>0\}.$$

This is well defined by our assumption that the path eventually escapes the bound. Since the M matrix corresponding to $T=T(x_0,\varepsilon_1)$ is a P-matrix, by Corollary 1 from the paper, $x_t(x_0,\varepsilon_1)$ is the unique solution for which $i_t>0$ for all $t>T=T(x_0,\varepsilon_1)$.

Also, for any given initial state $x_0\in\mathbb{R}^n$, and initial shock $\varepsilon_1\in\mathcal{E}$, we write $i_{T,t}(x_0,\varepsilon_1)$ and $x_{T,t}(x_0,\varepsilon_1)$ for the unique values of i_t and x_t, respectively, under perfect foresight given those initial conditions, with $\varepsilon_t = 0$ for $t > 1$, and where the bound is only imposed for the first T periods. (Uniqueness follows from Corollary 1 by the fact that the M-matrix corresponding to T is a P-matrix.) Now let:

$$\underline{i}_T(x_0, \varepsilon_1) := \min_{t>T} i_{T,t}(x_0, \varepsilon_1).$$

This is well defined as after period T, the endogenous variables x_t just follow a stationary VAR(1), for which there must exist a period $S > T$ such that for any x_0 and ε_1, $i_{T,S}(x_0, \varepsilon_1) > i_{T,T}(x_0, \varepsilon_1)$.

Now, since M is a P-matrix, the solution y of the associated LCP is Lipschitz continuous in the q vector (Mangasarian & Shiau 1987). Thus $\underline{i}_T$ must be continuous in x_0 and ε_1. Thus, for any $x_0 \in \mathbb{R}^n$, and $\varepsilon_1 \in \mathcal{E}$, $\underline{i}_{T(x_0,\varepsilon_1)}^{-1}((0,\infty))$ is open. Furthermore, by construction, $(x_0, \varepsilon_1) \in \underline{i}_{T(x_0,\varepsilon_1)}^{-1}((0,\infty))$, as $i_{T(x_0,\varepsilon_1),t}(x_0, \varepsilon_1) = i_t(x_0, \varepsilon_1)$ for all $t \in \mathbb{N}$, and $i_t(x_0, \varepsilon_1) > 0$ for $t > T(x_0, \varepsilon_1)$.

Consider the function $\phi: \mathbb{R}^n \times \mathcal{E} \to \mathbb{R}^n \times \mathcal{E}$ defined by:

$$\phi(x_0, \varepsilon_1) = (x_1(x_0, \varepsilon_1), 0),$$

for all $x_0 \in \mathbb{R}^n$, and $\varepsilon_1 \in \mathcal{E}$. $\mathbb{R}^n \times \mathcal{E}$ is a complete, proper metric space, as $\mathcal{E}$ is closed. Note that for all, for all $x_0 \in \mathbb{R}^n$, and $\varepsilon_1 \in \mathcal{E}$, on the open set $\underline{i}_{T(x_0,\varepsilon_1)}^{-1}((0,\infty)) \ni (x_0, \varepsilon_1)$, ϕ is Lipschitz continuous, as on this set it is sufficient to solve a size $T(x_0, \varepsilon_1)$ LCP and the M-matrix of this size is a P-matrix, which is sufficient for Lipschitz continuity (Mangasarian & Shiau 1987). Thus as ϕ is locally Lipschitz continuous everywhere, it is continuous everywhere. Additionally, by assumption ϕ has a unique fixed point $(x^*, 0)$, where x^* denotes the steady state of x_t, and for all $x_0 \in \mathbb{R}^n$, and $\varepsilon_1 \in \mathcal{E}$, $\phi^j(x_0, \varepsilon_1)$ converges to $(x^*, 0)$ as $j \to \infty$. Finally, there exists an open ball $U \ni x^*$ such that for any open set $V \ni x^*$, there exists $j \in \mathbb{N}$ such that $\phi^j(U) \subseteq V$, as for a small enough open ball U, ϕ is linear on U. This means that ϕ satisfies the conditions of the converse to Banach's fixed point theorem given in Daskalakis, Tzamos & Zampetakis (2018). Consequently, there exists a metric d^* on $\mathbb{R}^n \times \mathcal{E}$ which is topologically equivalent to the usual Euclidean one, and such that $(\mathbb{R}^n \times \mathcal{E}, d^*)$ is a complete metric space with:

$$d^*\big(\phi(x_0, \varepsilon_1), \phi(\tilde{x}_0, \tilde{\varepsilon}_1)\big) \le \frac{1}{2} d^*\big((x_0, \varepsilon_1), (\tilde{x}_0, \tilde{\varepsilon}_1)\big),$$

for all $x_0, \tilde{x}_0 \in \mathbb{R}^n$, and $\varepsilon_1, \tilde{\varepsilon}_1 \in \mathcal{E}$. Now $\tilde{\mathcal{X}} \times \mathcal{E}$ is closed and compact under the Euclidean metric, thus it is also closed and compact under d^*. Hence, $\tilde{\mathcal{X}} \times \mathcal{E}$ is bounded under the d^* metric, i.e., there exists $r > 0$ such that for all $(x_0, \varepsilon_1) \in \tilde{\mathcal{X}} \times \mathcal{E}$, $d^*\big((x_0, \varepsilon_1), (x^*, 0)\big) \le r$. Let:

$$\mathcal{W} := \{x_0 \in \mathbb{R}^n | \forall \varepsilon_1 \in \mathcal{E}, d^*\big((x_0, \varepsilon_1), (x^*, 0)\big) \le r\}$$

and:

$$\begin{aligned}\mathcal{X} &:= \{x_0 \in \mathbb{R}^n | \forall \varepsilon_1 \in \mathcal{E}, d^*\big((x_0, \varepsilon_1), (x^*, 0)\big) \le 2r\} \\ &= \bigcap_{\varepsilon_1 \in \mathcal{E}} \{x_0 \in \mathbb{R}^n | d^*\big((x_0, \varepsilon_1), (x^*, 0)\big) \le 3r\}\end{aligned}$$

Then $\tilde{\mathcal{X}} \subseteq \mathcal{W} \subseteq \mathcal{X}$, and $\mathcal{W}$ & $\mathcal{X}$ are closed sets, since they are intersections of closed sets. (The individual sets are closed since d^* is continuous under the d^* metric, so it is also continuous under the Euclidean metric.)

Furthermore:

$$\mathcal{X} \times \mathcal{E} \subseteq \{x_0 \in \mathbb{R}^n, \varepsilon_1 \in \mathcal{E} | d^*\big((x_0, \varepsilon_1), (x^*, 0)\big) \le 3r\},$$

which is a closed ball in $(\mathbb{R}^n \times \mathcal{E}, d^*)$, and hence is compact. Thus as $\mathcal{X} \times \mathcal{E}$ is a closed subset of a compact set, it is compact. Likewise, $\mathcal{W}$ is compact. Additionally, for all $x_0 \in \mathcal{X}$, $\varepsilon_1 \in \mathcal{E}$, $d^*\big(\phi(x_0, \varepsilon_1), (x^*, 0)\big) \le r$, so $x_1(x_0, \varepsilon_1) \in \mathcal{W} \subseteq \mathcal{X}$. Now let:

$$\mathcal{V} := \{(z, x_0) \in \mathbb{R}^n \times \mathcal{W} | d^*\big((x_0 + z, 0), (x_0, 0)\big) < r\},$$

$$\mathcal{U} := \{z \in \mathbb{R}^n | \forall (x_0, \varepsilon_1) \in \mathcal{W}, (z, x_0) \in \mathcal{V}\}.$$

$\mathcal{V}$ is open by the continuity of d^* and the equivalence of the metrics. Hence, by the compactness of $\mathcal{W}$, and the result of Theorem 2.3 of Escardó (2020), $\mathcal{U}$ is open as well. Thus, if we define $\mathcal{B}$ to be the closed unit ball in $\mathbb{R}^n$ under the Euclidean metric, there must exist some $\zeta > 0$ such that $\zeta\mathcal{B} \subseteq \mathcal{U}$. So, if $x_0 \in \mathcal{X}$, $\varepsilon_1 \in \mathcal{E}$, then for $z \in \zeta\mathcal{B} \subseteq \mathcal{U}$:

$$\begin{aligned}&d^*\big((x_1(x_0, \varepsilon_1) + z, 0), (x^*, 0)\big) \\ &\qquad \le d^*\big((x_1(x_0, \varepsilon_1) + z, 0), (x_1(x_0, \varepsilon_1), 0)\big) + d^*\big((x_1(x_0, \varepsilon_1), 0), (x^*, 0)\big) \\ &\qquad < r + r = 2r,\end{aligned}$$

as $x_1(x_0, \varepsilon_1) \in \mathcal{W}$. Thus, $x_1(x_0, \varepsilon_1) + \zeta\mathcal{B} \subseteq \mathcal{X}$ as required.

Now recall that for any $x_0 \in \mathbb{R}^n$, and $\varepsilon_1 \in \mathcal{E}$, $\underline{i}_{T(x_0,\varepsilon_1)}^{-1}((0, \infty))$ is an open set containing (x_0, ε_1). Consequently, $\{\underline{i}_{T(x_0,\varepsilon_1)}^{-1}((0, \infty)) | x_0 \in \mathcal{X}, \varepsilon_1 \in \mathcal{E}\}$ is an open cover of $\mathcal{X} \times \mathcal{E}$. Hence, by compactness, there exists finite sets $\hat{\mathcal{X}} \subseteq \mathcal{X}$ and $\hat{\mathcal{E}} \subseteq \mathcal{E}$ such that $\{\underline{i}^{-1}_{T(x_0,\varepsilon_1)}((0,\infty)) | x_0 \in \hat{\mathcal{X}}, \varepsilon_1 \in \hat{\mathcal{E}}\}$ is also an open cover of $\mathcal{X} \times \mathcal{E}$.

Now define $T^* := \max\{T(x_0,\varepsilon_1) | x_0 \in \hat{\mathcal{X}}, \varepsilon_1 \in \hat{\mathcal{E}}\}$. This is well defined as it is a maximum over a finite set. As the M corresponding to $T = T^*$ is a P-matrix, for all $x_0 \in \mathcal{X}$, and $\varepsilon_1 \in \mathcal{E}$, and $t \in \mathbb{N}$, $i_{T^*,t}(x_0,\varepsilon_1) = i_{T(x_0,\varepsilon_1),t}(x_0,\varepsilon_1) = i_t(x_0,\varepsilon_1)$ (again using Corollary 1). Thus, over the entire domain we only ever need to solve LCPs of size at most T^*. Finally, as the M corresponding to $T = T^*$ is a P-matrix, the LCP is Lipschitz in q, which is then sufficient for the Lipschitz conditions required for Theorem 3 (Restated) to apply. Hence, the model has a solution under rational expectations for shock distributions with sufficient mass at zero, completing the proof.

Appendix H.8: Proof of Proposition 12

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

$$\beta\sigma f^3 - \left((\alpha_{\Delta y} + \sigma)\beta + \gamma + \sigma\right) f^2 + \left((1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma\right) f - \alpha_{\Delta y} = 0. \tag{27}$$

When $f \leq 0$, 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 (27) 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:

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

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 (27) 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:

$$3 > f_1 + f_2 + f_3 = \frac{\big(\alpha_{\Delta y}+\sigma\big)\beta+\gamma+\sigma}{\beta\sigma},$$
$$3 > f_1f_2 + f_1f_3 + f_2f_3 = \frac{(1+\beta)\alpha_{\Delta y}+\gamma\alpha_\pi+\sigma}{\beta\sigma},$$
$$1 > f_1f_2f_3 = \frac{\alpha_{\Delta y}}{\beta\sigma},$$

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 (27) must be positive at $f = 1$, so:

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

Combining these inequalities gives the bounds:

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

Furthermore, if there are multiple solutions to equation (27), then the discriminant of its first derivative must be nonnegative, i.e.:

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

Therefore, we have the following bounds on α_π:

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

since,

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

as $\alpha_{\Delta y}<2\sigma-\frac{\gamma+\sigma}{\beta}$.

Consequently, there exists $\lambda,\mu,\kappa\in[0,1]$ such that:

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

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 (27)(27) 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 negative over the entire 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.:

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

Now, multiplying the denominator by f gives:

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

by equation (27). 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:

$$\begin{aligned}
&\frac{((1+\beta)\sigma+\gamma) - \sqrt{((1+\beta)\sigma+\gamma)^2 - 4\beta\sigma^2}}{2\beta\sigma} < f \\
&\qquad < \frac{((1+\beta)\sigma+\gamma) + \sqrt{((1+\beta)\sigma+\gamma)^2 - 4\beta\sigma^2}}{2\beta\sigma}.
\end{aligned}$$

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 (27) gives:

$$\begin{aligned}
&\Big[3\beta\sigma f^2 - 2\Big((\alpha_{\Delta y}+\sigma)\beta + \gamma + \sigma\Big)f + \Big((1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma\Big)\Big]\frac{\mathrm{d}f}{\mathrm{d}\alpha_{\Delta y}} = (1-\beta f)(1-f) \\
&\qquad > 0.
\end{aligned}$$

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

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

Note, however, that this expression is just the derivative of the left-hand side of equation (27) with respect to f.

To establish the sign of $\frac{df}{d\alpha_{\Delta y}}$, we consider two cases. First, suppose that equation (27) has three real solutions. Then, the unique solution to equation (27) in (0,1) is its lowest solution. Hence, this solution must be below the first local maximum of the left-hand side of equation (27). Consequently, at the $f \in (0,1)$, which solves equation (27), $3\beta\sigma f^2 - 2\Big((\alpha_{\Delta y}+\sigma)\beta + \gamma + \sigma\Big)f + \Big((1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma\Big) > 0$. Alternatively, suppose that equation (27) 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 $3\beta\sigma f^2 - 2\Big((\alpha_{\Delta y}+\sigma)\beta + \gamma + \sigma\Big)f + \Big((1+\beta)\alpha_{\Delta y} + \gamma\alpha_\pi + \sigma\Big) > 0$. 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 $(f, \alpha_{\Delta y})$ 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:

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

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

The former implies that:

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

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:

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

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

$$\alpha_{\Delta y} = \sigma\alpha_\pi,$$

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

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

Appendix H.9: Proof of Proposition 13

Defining $x_t = [x_{i,t} \quad x_{y,t} \quad x_{p,t}]'$, the price targeting model from 0 is in the form of Problem 2 (OBC), with:

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

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:

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

and so:

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

Now define:

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

so:

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

This is identical to the equation for f in Appendix H.8, 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 Appendix H.8, Assumption 1 holds for this model as well.

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

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

Appendix H.10: Proof of Proposition 14

If $\widetilde{X}$ is compact, then Γ is compact valued. Furthermore, $\widetilde{X}$ 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, $\widetilde{\mathcal{F}}$ is bounded above and below, so assumption 4.4 of SLP is satisfied as well. Furthermore, $\widetilde{\mathcal{F}}$ 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 $\mathcal{B}$ defined as in equation (28) and v^* defined as in equation (29)(29), $\mathcal{B}$ 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 $\mathcal{B}(v^*) = v^*$.

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

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

thus, our objective function is bounded above without additional assumptions. For a lower bound, we assume that for all $x \in \widetilde{X}$, $x \in \tilde{\Gamma}(x)$, 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 $\mathbb{V} := \{v | v: \tilde{\mathbb{X}} \to [-\infty, \infty)\}$ and $\underline{v}, \overline{v} \in \mathbb{V}$ by:

$$\underline{v}(x) = \frac{1}{1-\beta} \tilde{\mathcal{F}}(x_0, x_0),$$
$$\overline{v}(x) = \frac{1}{1-\beta} \left[u^{(0)} - \frac{1}{2} u^{(1)} \tilde{u}^{(2)^{-1}} u^{(1)'} \right],$$

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

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

$$\mathcal{B}(v)(x) = \sup_{z \in \tilde{\Gamma}(x)} \left[\tilde{\mathcal{F}}(x, z) + \beta v(z) \right] \tag{28}$$

for all $v \in \mathbb{V}$ and for all $x \in \tilde{\mathbb{X}}$. Then $\mathcal{B}(\underline{v}) \geq \underline{v}$ and $\mathcal{B}(\overline{v}) \leq \overline{v}$. Furthermore, if some sequence $(x_t)_{t=1}^{\infty}$ satisfies the constraint that for all $t \in \mathbb{N}^+$, $x_t \in \tilde{\Gamma}(x_{t-1})$, and the objective in (16) 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 $\mathcal{B}$ has a unique fixed point in $[\underline{v}, \overline{v}]$ to which $\mathcal{B}^k(\underline{v})$ converges pointwise, monotonically, as $k \to \infty$, and which is equal to the function $v^*: \tilde{\mathbb{X}} \to \mathbb{R}$ defined by:

$$v^*(x_0) = \sup\left\{ \textstyle\sum_{t=1}^{\infty} \beta^{t-1} \tilde{\mathcal{F}}(x_{t-1}, x_t) \,\middle|\, \forall t \in \mathbb{N}^+, x_t \in \Gamma(x_{t-1}) \right\}, \tag{29}$$

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

Furthermore, if we define:

$$\mathbb{W} := \{v \in V | v \text{ is continuous on } \tilde{\mathbb{X}}, v \text{ is concave on } \tilde{\mathbb{X}}\},$$

then as $\tilde{u}^{(2)}$ is negative-definite, $\underline{v} \in \mathbb{W}$. Additionally, under the assumption that $\tilde{\Gamma}(x)$ is compact valued, if $v \in \mathbb{W}$, then $\mathcal{B}(v) \in \mathbb{W}$, by the Theorem of the Maximum,[48] and, furthermore, there is a unique policy function which attains the supremum in the definition of $\mathcal{B}(v)$. Moreover, $v^* = \lim_{k \to \infty} \mathcal{B}^k(\underline{v})$ is concave and lower semi-continuous on $\tilde{\mathbb{X}}$.[49] We just need to prove that v^* is upper semi-continuous.[50] Suppose for a contradiction that it is not, so there exists $x^* \in \tilde{\mathbb{X}}$ such that:

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

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

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

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

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

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

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

$$v^*\big(x_0^{(\epsilon)}\big) < \delta + \sum_{t=1}^\infty \beta^{t-1}\tilde{\mathcal{F}}\big(x_{t-1}^{(\epsilon)}, x_t^{(\epsilon)}\big).$$

Hence:

$$\sum_{t=1}^\infty \beta^{t-1}\tilde{\mathcal{F}}\big(x_{t-1}^{(\epsilon)}, x_t^{(\epsilon)}\big) > v^*\big(x_0^{(\epsilon)}\big) - \delta > v^*(x^*).$$

Now, let $\mathcal{S}_0 := \{x \in \widetilde{X} \,|\, \|x^* - x\|_\infty \le 1\}$, and for $t \in \mathbb{N}^+$, let $\mathcal{S}_t := \Gamma(\mathcal{S}_{t-1})$. Then, since we are assuming Γ is compact valued, for all $t \in \mathbb{N}$, $\mathcal{S}_t$ is compact by the continuity of Γ. Furthermore, for all $t \in \mathbb{N}$ and $\epsilon \in (0,1)$, $x_t^{(\epsilon)} \in \mathcal{S}_t$. Hence, $\prod_{t=0}^\infty \mathcal{S}_t$ 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 $x_t := \lim_{k\to\infty} x_t^{(\epsilon_k)}$, and note that $x^* = x_0 \in \mathcal{S}_0 \subseteq \widetilde{X}$, and that for all $t, k \in \mathbb{N}^+$, $x_t^{(\epsilon_k)} \in \Gamma\Big(x_{t-1}^{(\epsilon_k)}\Big)$, so by the continuity of Γ, $x_t \in \Gamma(x_{t-1})$ for all $t \in \mathbb{N}^+$. Thus, by Fatou's Lemma:

$$v^*(x^*) \ge \sum_{t=1}^\infty \beta^{t-1}\tilde{\mathcal{F}}(x_{t-1}, x_t) \ge \limsup_{k\to\infty} \sum_{t=1}^\infty \beta^{t-1}\tilde{\mathcal{F}}\big(x_{t-1}^{(\epsilon,k)}, x_t^{(\epsilon,k)}\big) > v^*(x^*),$$

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

Appendix H.11: Proof of Proposition 15

Suppose that $(x_t)_{t=1}^\infty$, $(\lambda_t)_{t=1}^\infty$ satisfy the KKT conditions given in equations (18) and (19), and that $x_t \to \mu$ and $\lambda_t \to \bar{\lambda}$ as $t \to \infty$. Let $(z_t)_{t=0}^\infty$ satisfy $z_0 = x_0$ and $z_t \in \tilde{\Gamma}(z_{t-1})$ for all $t \in \mathbb{N}^+$. Then, by the KKT conditions and the concavity of:

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

we have that for all $T \in \mathbb{N}^+$:[51]

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

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

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

Thus:

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

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 $\lim_{T \to \infty} \beta^T z_T$:

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

which implies that $(x_t)_{t=1}^{\infty}$ solves Problem 4 (Linear-Quadratic).

References

Adjemian, Stéphane, Houtan Bastani, Fréderic Karamé, Michel Juillard, Junior Maih, Ferhat Mihoubi, George Perendia, Johannes Pfeifer, Marco Ratto & Sébastien Villemot. 2011. *Dynare: Reference Manual Version 4*. CEPREMAP.

Adjemian, Stéphane, Matthieu Darracq Pariès & Stéphane Moyen. 2007. *Optimal Monetary Policy in an Estimated DSGE for the Euro Area*. European Central Bank.

Aliprantis, C.D. & 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.

Ascari, Guido & Tiziano Ropele. 2009. 'Trend Inflation, Taylor Principle, and Indeterminacy'. *Journal of Money, Credit and Banking* 41 (8): 1557–1584.

Bai, Zhong-Zhi & D.J. Evans. 1997. 'Matrix Multisplitting Relaxation Methods for Linear Complementarity Problems'. *International Journal of Computer Mathematics* 63 (3–4) (January 1): 309–326.

Benigno, Pierpaolo & 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 & Charles M. Kahn. 1980. 'The Solution of Linear Difference Models under Rational Expectations'. *Econometrica* 48 (5): 1305–1311.

Boneva, Lena Mareen, R. Anton Braun & Yuichiro Waki. 2016. 'Some Unpleasant Properties of Loglinearized Solutions When the Nominal Rate Is Zero'. *Journal of Monetary Economics* 84 (December 1): 216–232.

Brendon, Charles, Matthias Paustian & Tony Yates. 2013. 'The Pitfalls of Speed-Limit Interest Rate Rules at the Zero Lower Bound'. Bank of England.

———. 2016. 'Self-Fulfilling Recessions at the Zero Lower Bound'.

———. 2019. 'Self-Fulfilling Recessions at the Zero Lower Bound'. *Journal of Monetary Economics* (June 15).

Cahill, Nathan D., John R. D'Errico, Darren A. Narayan & Jack Y. Narayan. 2002. 'Fibonacci Determinants'. *The College Mathematics Journal* 33 (3) (May 1): 221–225.

Carlstrom, Charles T., Timothy S. Fuerst & Matthias Paustian. 2015. ‘Inflation and Output in New Keynesian Models with a Transient Interest Rate Peg’. *Journal of Monetary Economics* 76 (November): 230–243.

Cottle, Richard W., Jong-Shi Pang & 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.

Cúrdia, Vasco, Marco Del Negro & Daniel L. Greenwald. 2014. ‘Rare Shocks, Great Recessions’. *Journal of Applied Econometrics* 29 (7): 1031–1052.

Daskalakis, Constantinos, Christos Tzamos & Manolis Zampetakis. 2018. ‘A Converse to Banach’s Fixed Point Theorem and Its CLS-Completeness’. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, 44–50. STOC 2018. New York, NY, USA: Association for Computing Machinery.

Del Negro, Marco, Marc Giannoni & Christina Patterson. 2015. *The Forward Guidance Puzzle*. Staff Reports. Federal Reserve Bank of New York.

Dontchev, Asen L. & R. Tyrrell Rockafellar. 2014. *Implicit Functions and Solution Mappings: A View from Variational Analysis*. Springer.

Escardó, Martín Hötzel. 2020. ‘Intersections of Compactly Many Open Sets Are Open’. *ArXiv:2001.06050 [Math]* (January 16).

Fagan, Gabriel, Jerome Henry & Ricardo Mestre. 2005. ‘An Area-Wide Model for the Euro Area’. *Economic Modelling* 22 (1): 39–59.

Fair, Ray C & John B Taylor. 1983. ‘Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models’. *Econometrica* 51 (4) (July): 1169–85.

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

Farmer, Roger E.A., Vadim Khramov & Giovanni Nicolò. 2015. ‘Solving and Estimating Indeterminate DSGE Models’. *Journal of Economic Dynamics and Control* 54 (May 1): 17–36.

Fernández-Villaverde, Jesús, Grey Gordon, Pablo Guerrón-Quintana & Juan F. Rubio-Ramírez. 2015. 'Nonlinear Adventures at the Zero Lower Bound'. *Journal of Economic Dynamics and Control* 57 (C): 182–204.

Guerrieri, Luca & Matteo Iacoviello. 2015. 'OccBin: A Toolkit for Solving Dynamic Models with Occasionally Binding Constraints Easily'. *Journal of Monetary Economics* 70 (March): 22–38.

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

———. 2016. 'Computation of Solutions to Dynamic Models with Occasionally Binding Constraints.' Unpublished.

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

Jiang, Shunhua, Zhao Song, Omri Weinstein & Hengjie Zhang. 2020. 'Faster Dynamic Matrix Inverse for Faster LPs'. *ArXiv:2004.07470 [Cs]* (April 16).

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 & 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 & 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.

Li, Cui-Xia & Shi-Liang Wu. 2016. 'A Note on the Unique Solution of Linear Complementarity Problem'. *Cogent Mathematics* 3 (1) (December 20).

Mangasarian, O. L. & T.-H. Shiau. 1987. 'Lipschitz Continuity of Solutions of Linear Inequalities, Programs and Complementarity Problems'. *SIAM Journal on Control and Optimization* 25 (3) (May 1): 583–595.

McKay, Alisdair, Emi Nakamura & Jón Steinsson. 2017. 'The Discounted Euler Equation: A Note'. *Economica* 84 (336): 820–831.

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

Ruiz, Daniel. 2001. *A Scaling Algorithm to Equilibrate Both Rows and Columns Norms in Matrices*. CM-P00040415.

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

Smets, Frank & 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 & E. Prescott. 1989. *Recursive Methods in Economic Dynamics*. Harvard University Press.

Swarbrick, Jonathan, Tom D. Holden & Paul Levine. 2016. 'Endogenous Financial Crises in a Model of Occasionally Binding Constrained Borrowing'.

Trefethen, L.N. & 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.

Wright, Thomas G & 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.