Adaptive Learning and Cognitive Discounting: Can Depressions be Stable?*

George W. Evans, University of Oregon and University of St Andrews
Seppo Honkapohja, Aalto University School of Business
Kaushik Mitra, University of Birmingham
Bruce McGough, University of Oregon

March 23, 2026

Preliminary and Incomplete†

Abstract

Nonlinear New Keynesian models can have multiple REE steady states that are locally stable under adaptive learning (AL). Gabaix (2020) introduced cognitive discounting (CD) as a behavioral modification with the potential to resolve some well-known policy puzzles in long-horizon models. We argue that boundedly rational agents would naturally combine AL and CD, leading to novel results. In our framework this can yield a model with two locally stable steady states: the one targeted by policy, and a 'Great Depression' steady state. We explore the dynamics and policy implications of this framework. An unlucky series of shocks can push the economy away from the target steady state and into the domain of attraction of the Depression steady state. We examine which policies can dislodge the economy from the Depression steady state and return it to the target steady state.

*The second author is grateful for support from the Yrjö Jahnsson Foundation. A draft of this paper was presented at the "Expectations in Dynamic Macroeconomic Models" Workshop held at UNSW Sydney, Australia, Dec. 18, 2025. We are grateful for comments at the Workshop and in earlier presentations to the University of Oregon Macroecomoics group and at a University of St Andrews Economics brownbag talk.

†Please do not quote or cite without permission

1 Introduction

In Evans, Honkapohja, and Mitra (2022) (EHM) we developed a nonlinear New Keynesian model with three rational expectation (RE) steady states, and showed that the steady state targeted by monetary policy is locally, but not globally, stable under adaptive learning. In particular, sufficiently large adverse shocks leading to depressed expectations for future aggregate output could lead to paths entering a stagnation region unless offset by suitable monetary and fiscal polices. The adaptive learning (AL) approach to household and firm expectations of future aggregate state variables (e.g. GDP and inflation) is to model them as econometric forecasts, based on data available to agents, with forecasting equation coefficients updated over time in line with new observations. AL is attractive because, while boundedly rational, it has the potential to converge over time to RE, and also because it can be viewed as a corresponding to how economists and policymakers make forecasts.

The current paper extends our framework in several ways. A major contribution of the paper is to incorporate cognitive discounting into AL. Cognitive discounting (CD) was introduced by Gabaix (2020) within an RE framework: while anchored to RE, forecasts for dates further into the future were correspondingly discounted toward the RE steady state. The motivation given in Gabaix (2020) was behavioral. However, there is a separate motivation in the adaptive learning approach: longer-horizon econometric forecasts can be subject to increased forecast uncertainty due to non-stationarity of underlying stochastic processes, including the possibility of future structural change.

In this paper we show how to formally integrate geometric discounting of forecasts into AL using a multiplicative modelling approach. This provides an efficient framework for studying the global dynamics of output and inflation expectations, and the implied time-paths of realized output and inflation. In addition our approach facilitates computations of IRFs (impulse response functions) starting from any steady state (or, indeed, from any initial state).

An additional advantage of our approach is that it can readily incorporate a variety of exogenous stochastic shocks. In the current paper we focus on exogenous productivity and mark-up shocks, but in work in progress we have been able to incorporate discount factor, sales tax, financial, and labor supply shocks.

1.1 Nonlinear NK model with Adaptive Learning and Cognitive Discounting

Gabaix (2020) introduced "Cognitive Discounting" (CD) as a behavioral modification to rational expectations (RE) within an NK model:

"The main nonstandard feature of the model is a form of cognitive discounting. The phenomenon I

want to analyze is that the world is not fully understood by agents, especially events that are far into the future. To capture this mathematically, I make the following key assumption: agents, as they simulate the future, shrink their simulations toward a simple benchmark, namely, the steady state of the economy."

To illustrate, Gabaix implements this in the canonical 2-equation linearized NK model as

$$\begin{aligned} \text{IS:} \quad & x_t = ME_t x_{t+1} - \sigma\left(i_t - E_t\pi_{t+1} - r_t^n\right) \\ \text{PC:} \quad & \pi_t = M^f E_t \pi_{t+1} + \kappa x_t. \end{aligned}$$

where $M, M^f \in [0,1]$ are "attention parameters" of consumers and firms, respectively, capturing their partial myopia toward distant events. These shrink forecasts in the direction of the steady state, which in the above equations is normalized to zero. $M = M^f = 1$ correspond to RE. Gabaix shows that when RE is modified in this way, the properties of the NK model are substantially changed in a number of significant aspects (e.g. there is no forward guidance puzzle.)

Before turning to our formal model, we briefly discuss the underlying reasoning of the AL and CD approaches. While the motivation given in Gabaix (2020) was behavioral, we believe there is a distinct motivation within the AL approach, which is that time-series forecasts of output and inflation further into the future can have increasing forecast uncertainty due to perceived non-stationarity of exogenous stochastic processes and especially to the possibility of future structural change. In our setting, there is a steady state targeted by policy, and we assume that, for forecasts further into the future, agents modify their econometric forecasts by placing a correspondingly higher weight on these policy targets.

The AL critique of REE is that REE needs a rationale for how it would be attained by agents because: (i) agents are unlikely to know all the structural parameters of the true model, and (ii) even if they did, it is not optimal to have REE expectations unless they confidently believe all other agents have these expectations (REE is in essence a Nash equilibrium) and (iii) in many models there are multiple REE. A particular REE then requires a coordination mechanism (which AL can provide for "expectationally stable" REE).

A major attraction of the AL approach is that, given the statistical forecasting model used by agents (which is usually assumed to nest or approximate the form of an RE forecasting rule, but with estimated coefficients that are updated over time), AL provides a way of way to calculate recursively the evolution over time of both the agent's expectations (including the time profile of their forecasts of future variables) and the corresponding realized path of the economy. In the present paper the key variables of interest are aggregate output y_t and inflation π_t, and agents' forecasts of their future values $\left(y_{t+s}^e, \pi_{t+s}^e\right)$.

From the AL perspective, the CD approach to modifying RE forecasts remains subject to the AL critique of REE, because the CD framework in Gabaix (2020) requires that agents *first* compute RE forecasts and *then*, having done that, discount RE forecasts in the direction of "the" steady state. Thus, it assumes that agents first embrace RE and then move away

from it.

However, an attractive approach is one that combines **both** AL **and** CD. Specifically, households and firms make decisions *starting with AL forecasts* of future aggregate output and inflation rates. AL is a plausible model of rationality because economists in practice make forecasts using time-series econometric models. However, econometric forecasts over longer horizons are subject to unpredictable structural change. The statistical structure of the economy includes nonstationary processes that renders long-range forecasts to be less reliable, e.g. see Hendry and Nielsen (2007), Hendry and Pretis (2016)) and Castle, Clements, and Hendry (2019). Due to the reduced confidence agents should have in long-range econometric forecasts, those forecasts should be correspondingly discounted. Because policymakers have clearly stated goals, which we interpret to be full-employment output and inflation at the Central Bank target of 2% per year, it is natural for longer-horizon forecasts to be discounted toward the steady state around which policymakers aim to stabilize the economy.

1.2 Structure of the Paper and the Main Results

In Section 2 we summarize the nonlinear NK model, based on Benhabib, Evans, and Honkapohja (2014) and extended in Evans, Honkapohja, and Mitra (2022) (EHM). The EHM model has 3 steady states: the steady state targeted by peacemakers is locally stable under AL; there is a steady state that is induced by the ZLB, but which is unstable under AL; and there is also a very low-level steady state, underpinned by government spending, which is also locally stable under learning.

Section 2.2 introduces adaptive learning, describes how AL forecast parameters are updated over time, in both nonstochastic and stochastic setups. AL follows a "temporary equilibrium" approach, in which current decisions by agents depend on their forecasts of future inflation and output, as well as exogenous variables. Then, in the next period, expectations are revised in light of the most recent outcomes, and the equilibrium time path evolves recursively..

For the simplest setting, in which the model is nonstochastic, or there are only *iid* random shocks, the key results can be presented using a phase diagram in expectations space (π^e, y^e), showing the direction of motion of the expectations that results from the corresponding realized (π, y) and the AL updating rule for expectations. When exogenous random shocks are included, it is straightforward to simulate time series paths for (π_t, y_t). In this case the "Perceived Law of Motion" of the forecasting rules for $\left(y^e_{t+s}, \pi^e_{t+s}\right)$ are a function of current exogenous shocks, with coefficients that are updated recursively over time, and realized actual output and inflation (π_t, y_t) depend on the whole forecast profile $\left(y^e_{t+s}, \pi^e_{t+s}\right)$ as well as on the current exogenous shocks. This recursive structure generates the time path of the economy.

Section 3 then describes how forecasts of future output, and inflation and real interest rates are modified by also incorporating CD of expectations of future output, inflation and real interest rates: AL expectations are discounted geometrically toward the steady state values targeted by policymakers. The key equations use a multiplicative set-up that makes use of results on Mahler series, which can provide code for rapid numerical calculation of phase diagrams, impulse response functions and policy evaluation.

The dynamics of the model under AL and CD are again most conveniently presented using phase diagrams showing the direction of change over time in expected inflation and expected output (π^e, y^e). The qualitative features of the phase diagrams depend, of course, on both structural parameters and CD parameters. For the structural parameters we in most cases follow the values used in Evans, Honkapohja, and Mitra (2022). These are provided in Table 1 and discussed further in the Appendix. Apart from the inclusion of CD parameters, the main change is to the Rotemberg-type price-adjustment cost function, in which the tails are modified to ensure that inflation rates at the high end, and deflation rates at the low end, are consistent with historical data.

1.3 Depressions

The qualitative features of the model, not surprisingly, depend on the CD parameters chosen. Keeping the values of the non-CD structural parameters unchanged and varying the CD parameters numerically, we find two broad cases. If the CD parameters $0 \leq m_y, m_r \leq 1$ are small enough, then there is a unique steady state at the target steady state, as we would expect, because future expectations of output and inflation are discounted heavily toward the target steady state. However, for a range of CD parameters (m_y, m_r) with values closer to 1, so that the discounting of AL expectations is less severe, there continue to be 3 steady states, but now the low output/low inflation steady state consistent with a depressed economy. To understand the implications of this, in Section 4.1 we choose a calibration of (m_y, m_r) that gives a depression steady state with output and deflation values close in size to the trough of the US Great Depression of the early 1930s. The corresponding phase diagram shows that with this calibration, both the target steady state and the Depression steady state are locally stable. For some initial expectations (π^e, y^e) there will be extreme sensitivity of the path over time to the precise initial conditions (and to the stochastic shocks realized).

These results raise two major questions: (i) Starting from the target steady state, is it possible that an adverse series of shocks can push the economy into the depression steady state? Using simulations we show that an unlucky series of productivity and mark-up shocks can indeed lead to and trap the economy in depression. (ii) If the economy is in the depression steady state, are there policies that would dislodge the economy from depression and return it to the target steady state? Our answer to this question is also yes, though we will show that the policies need to be carefully chosen.

Eggertsson (2008) and Eggertsson (2012) discussed the New Deal policies for overcoming the very low output accompanied by deflation that was present in the US economy in the early 1930s. The explanation in Eggertsson (2008) focused on Roosevelt and the New Deal providing a "policy regime change" that generates an endogenous shift in expectations from "contractionary" (expected future economic contraction and deflation) to "expansionary" ("expected future economic expansion and inflation"). The focus of Eggertsson (2012) is the NIRA policies to increase prices and wages to reduce or eliminate deflation.

The major contribution of the present paper is our analysis of the *dynamics* of output and inflation expectations over time, and the implications of expectational learning dynamics for whether specific policies, in an economy now far from target due to adverse shocks, can successfully return the economy to the target steady state. Using our benchmark calibration we obtain the respective domains of attraction (DOAs) of the target and depression steady states. This provides crucial information on how far expectations can be from the target steady state without following a path to Depression.

In our baseline model we also include, as is standard in New Keynesian models, serially correlated exogenous productivity and mark-up shocks. To understand their effects, in Section 5 we use our model to compute "net IRFs", the effects over time, starting from the target steady state, on realized output and inflation, and on output and inflation expectations, of productivity and mark-up (elasticity of substitution) shocks. This section also illustrates the possibility of "escapes": adverse sequences of shocks that push the economy out of the target steady state DOA and into the Depression DOA. Related to this, in our baseline calibration the DOA of the depression steady state includes expectations (y^e, π^e) associated with moderate recessions and deflation rates. In particular, an expected output gap of 6.2% below target output accompanied by expected deflation of 5.0% lies just inside the DOA of the Depression steady state.

1.4 Policy

In Sections 6, 7 and 8, we assume the economy is in the depression steady state, and consider three types of policy for pushing the economy out of depression and back to the target steady state. To study this we use the AL framework, introduced in Evans, Honkapohja, and Mitra (2009), in which policymakers announce and credibly commit to a temporary policy, e.g. to increase government spending for T_p periods, after which government spending returns to its previous level. Agents alter their decision rules to immediately take into account the direct effects of the policy (in this case the corresponding increase in taxes). Then, as agents move forward in time, they observe the time paths of y and π, which are used to revise their expectations (y^e, π^e) and so capture indirect effects of the policy as the economy responds to the policy over time.

We first consider a large increase in government spending for a fixed period of time. Using

simulations we find that this fails to dislodge the economy from the DOA of the depression steady state: following the policy, the economy sinks back into Depression. The reason for its lack of success is the well-known effect from RBC models, that due to the income effect, there is a negative effect on inflation, worsening the deflation associated with the Depression steady state. We next study policies to increase the price level. We model this as an announced and credibly implemented increase in the mark-up or a fixed period of time. In line withEggertsson (2012), this policy, for a large enough increase in the mark-up for a suitable length of time, is successful in moving the economy into the DOA of the target steady state. Finally, we return to government spending but consider a variation in which during the policy period, the increased government spending is on goods which are strongly substitutable with between private consumption. For a large enough increase in government spending of this type for a suitable length in time, the policy is also successful. This is because the negative income effect emphasized in RBC models is negligible for this type of government spending.

2 Model Structure: Nonlinear New Keynesian model

We extend the global nonlinear New Keynesian (NK) model with adaptive learning, introduced and developed in Benhabib, Evans, and Honkapohja (2014) and Evans, Honkapohja, and Mitra (2022) (EHM), to incorporate cognitive discounting within the learning framework. As shown in EHM this model has three steady states: the steady state targeted by policymakers, the well-known unintended steady state with below target inflation, induced by the ZLB, and a third steady state at very low output $y_t = g$ accompanied by strong deflation.

2.1 The Nonlinear Model

In the nonlinear model of EHM there is a continuum of household-producers i that maximize utility subject to flow budget and production constraints

$$E_t \sum_{t=0}^{\infty} \beta^t \left\{ \log\left(c_{t,i} + \xi g_t\right) - (1+\varepsilon)^{-1} h_{t,i}^{1+\varepsilon} - \phi\left(\frac{P_{t,i}}{P_{t-1,i}} - \pi^*\right)\right\},$$

$$\text{s.t. } c_{t,i} + m_{t,i} + b_{t,i} + \Upsilon_{t,i} = m_{t-1,i}\pi_t^{-1} + R_{t-1}\pi_t^{-1} b_{t-1,i} + \frac{P_{t,i}}{P_t} y_{t,i},$$

for discount factor $0 < \beta < 1$.

Household production is given by $y_{t,i} = A_t h_{t,i}^a$, where $0 < \alpha < 1$ and $A_t > 0$ is a stationary process. Here $P_{t,i}$ denotes price of consumption good i at time t, $c_{t,i}$ denotes consumption of good i at time t, and $h_{t,i}$ denotes labor input into production of good variety i. For

convenience we assume the cashless limit of the economy. Here $b_{t,i}$ is end-of-period bond holding and $\Upsilon_{t,i} = g_t$ is a lump-sum tax. $\phi \geq 0$ captures a Rotemberg-type convex pricing friction with $\phi(\pi^*) = 0$, where π^* is the policy inflation target, and $\pi_t = P_t/P_{t-1}$ denotes the time t inflation factor, where P_t is the aggregate price level at t. The form of $\phi(\pi)$ is based on the Linex function, which can incorporate asymmetry of the adjustment cost around π^*, see (**??**) below for details.

Utility of consumption includes both private consumption $c_{t,i}$ and public consumption $g_t > 0$ of the goods aggregator. The consumption aggregator takes the usual constant elasticity of substitution (CES) form, with elasticity of substitution between two goods $\nu_t > 1$, where ν_t is either constant or a stationary AR(1) process. Output is differentiated and operates under monopolistic competition, with downward-sloping demand curve $P_{t,i} = (y_{t,i}/y_t)^{-1/\nu_t} P_t$, where $P_t = \left[\int_0^1 P_{t,i}^{1-\nu_t} di\right]^{1/(1-\nu_t)}$.

The parameter $0 < \xi \leq 1$ captures the degree of substitution between private and public consumption, as in Christiano and Eichenbaum (1992). Except when examining fiscal policy we will assume $g_t = g > 0$ is constant over time. We consider the representative-agent equilibrium in which all agents behave identically and receive the same amount of the government consumption good. The government satisfies the usual flow budget constraint and households are Ricardian. Households must pay taxes and this places a lower bound on output $y_{t,i} \geq g$.

We follow the long-horizon approach introduced by Preston (2005). Our boundedly-rational agents are assumed to be a version of long-horizon anticipated-utility optimizers, as in Eusepi (2010), Eusepi and Preston (2018).[1] In EHM this gave the temporary equilibrium equations for aggregate output and inflation, conditional on agent forecasts. The **temporary equilibrium equation for output** is

$$y_t = \max\left\{g_t, (1-\xi)\, g_t \sum_{s=1}^{\infty} \hat{D}^e_{t,t+s}\left(y^e_{t+s} - (1-\xi)\, g^e_{t+s}\right)\right\}. \tag{1}$$

Here $\hat{D}^e_{t,t+s} = \left(\prod_{j=1}^{s} r^e_{t+j}\right)^{-1}$, where $r^e_{t+j} = R_t = R\left(\pi^e_{t+1}, y^e_{t+1}\right)/\pi^e_{t+j}$, is the expected discount factor with (known) nominal interest rate policy rule given by

$$R_t = R(\pi^e_{t+1}, y^e_{t+1}) = 1 + (R^* - 1)\left(\pi^e_{t+1}/\pi^*\right)^{BR^*/(R^*-1)}\left(y^e_{t+1}/y^*\right)^{\phi_y}, \tag{2}$$

where $B > 1$ and $\varphi_y \geq 0$ are parameters of the policy rule. Here $R^* = \beta^{-1}\pi^*$ is the target nominal interest-rate factor at π^*, and y^* is the nonstochastic steady state output corresponding to the policy inflation target $\pi = \pi^* \geq 1$. By design, the interest-rate rule satisfies

[1] However, for the reasons given in Benhabib, Evans, and Honkapohja (2014) and Evans, Honkapohja, and Mitra (2022),we agents impose the observation that in temporary equilibrium, at every moment in time t, $P_{t,i}/P_t = 1$.

the ZLB, but for $\pi^* > 1$ policy will set R_t arbitrarily close to 1 for low sufficiently π^e_{t+1} and y^e_{t+1}.

The temporary equilibrium output equation (1) is derived under the assumption that agents have point expectations, a bounded rationality assumption that we find most plausible given the nonlinearities present in the model. Thus y_t depends on g_t and $\left\{y^e_{t+s}, \pi^e_{t+s}, g^e_{t+s}\right\}_{s\geqslant 1}$. Equation (1) is a standard Aggregate Demand (AD or IS curve) equation based on the consumption Euler equation and the intertemporal budget constraint coming from discounted expected future incomes and taxes.

The **temporary equilibrium equation for inflation** is

$$Q(\pi_t) = G(y_t) + \sum_{s=1}^{\infty} \beta^s G(y^e_{t+s}), \text{ where } Q(\pi) = \phi'(\pi)\pi \text{ and} \tag{3}$$

$$G(y_t) = \frac{\nu_t}{\alpha}\left(\frac{y_t}{A_t}\right)^{(1+\varepsilon)/\alpha} - \frac{(\nu_t - 1)y_t}{y_t - (1-\xi)g}, \tag{4}$$

where Q is an (invertible) increasing function for the class of price-adjustment cost functions $\phi(\pi)$ that we consider. Note that while the AD curve depends on both expected future outputs and inflation rates, the AS curve then depends only on current and expected future outputs. As discussed below, if A_t and ν_t are stochastic, e.g. AR(1) processes, then $G(y^e_{t+s})$ will depend on A^e_{t+s} and ν^e_{t+s} as well as y^e_{t+s}, but we will use the bounded rationality assumption of point expectations.

To interpret the AS (Phillips-curve) equation (14), it is useful to split $G(y)$ into its two pieces:

$$Q(\pi_t) = G(y) + \sum_{n\geq 1} \beta^n G(y^e_n), \text{ where}$$

$$G(y) = \underbrace{\frac{\nu}{\alpha}\left(\frac{y}{A}\right)^{\frac{1+\varepsilon}{\alpha}}}_{\text{marginal cost}} - \underbrace{\frac{y(\nu-1)}{y-(1-\xi)g}}_{\text{income effect}}.$$

Higher expected output implies higher marginal cost and thus higher π. However, higher g also has an income effect because it leads to lower after-tax real income, which reduces consumption of all goods, including leisure. The consequent increased labor supply lowers wages and thus lowers π. This is an effect well-known from real business cycle and related "Classical" models, but is also, as we will show, present in our New Keynesian framework.

2.2 Adaptive Learning

Following the AL approach, we assume agents use log-linear forecast rules for future y and π that are updated using recursive LS. The nonstochastic case with $A_t = A, \nu_t = \nu, g_t = \bar{g}$

is simplest. Under "steady-state" learning

$$\begin{aligned} y^e_{t+s} &= y^e_t, \text{ and } \pi^e_{t+s} = \pi^e_t \text{ for all } s > 0, \text{ where} \\ y^e_t &= y^e_{t-1} + \omega(y_{t-1} - y^e_{t-1}) \text{ and } \pi^e_t = \pi^e_{t-1} + \omega(\pi_{t-1} - \pi^e_{t-1}). \end{aligned}$$

Here $0 < \omega < 1$ is the learning "gain" parameter, usually assumed "small."

The TE (temporary equilibrium) equations then simplify to a nonlinear system

$$y_t = T_y(\pi^e_t, y^e_t, g),\ \pi_t = S(\pi^e_t, y^e_t, g).$$

There are **three steady states,** two of which are stable under AL: the targeted (full-employment) steady state $\bar{y}$, and a subsistence steady state $y_t = g$.

The AL approach can be extended to stochastic economies. For example, it is straightforward to include exogenous stationary AR(1) shocks to productivity $A_t/\bar{A}$ and the mark-up $\nu_t/\bar{\nu}$. There are alternative ways to model the AL forecasts that take account of these shocks, but in practice results are not sensitive to the method chosen, and local stability of each steady state is unaffected by the shocks. If agents observe the shocks A_t and ν_t and know the means and AR1 parameters, they can suitably adjust their AL forecasts y^e_{t+s} and π^e_{t+s}.

Alternatively agents could use recursive least squares (RLS) to directly estimate the PLM

$$\begin{aligned} \ln(y_t) &= f_y + d_{yA}\ln(\tilde{A}_t) + d_{y\nu}\ln(\tilde{\nu}_t) + \eta_{yt} \\ \ln(\pi_t) &= f_\pi + d_{\pi A}\ln(\tilde{A}_t) + d_{\pi\nu}\ln(\tilde{\nu}_t) + \eta_{\pi t}, \end{aligned}$$

where $\tilde{A}_t = A_t/\bar{A}$ and $\tilde{\nu}_t = \nu_t/\bar{\nu}$ and to use current parameter estimates and the known AR parameters to iterate forward to obtain $t+s$ forecasts. If AR(1) parameters are unknown agents can obtain and use consistent estimates of them. In practice "constant gain" RLS is typically used, in which a small weight $0 < \omega < 1$ is placed on each new observation.

Allowing for stochastic exogenous AR(1) shocks to A_t and ν_t, if agents use point expectations then the AS equation is given by (14) with

$$G\left(y^e_{t+s}, A^e_{t+s}, \nu^e_{t+s}\right) = \frac{\nu^e_{t+s}}{\alpha}\left(\frac{y^e_{t+s}}{A^e_{t+s}}\right)^{(1+\varepsilon)/\alpha} - \frac{\left(\nu^e_{t+s} - 1\right) y^e_{t+s}}{y^e_{t+s} - (1-\xi)\, g}.$$

When exogenous stochastic shocks like these are included, rational expectation steady states are stochastic processes. However, local stability results under AL are essentially unchanged if the support of the exogenous shocks is sufficiently small. The dynamic path under RLS adaptive learning is specified recursively as before. Under AL convergence to an E-stable REE will be to a *Restricted Perceptions Equilibrium* (RPE) that is "near" an REE, rather than to the REE itself, because of the use of steady-state learning and/or the use of linear forecast models within a nonlinear model.

In the AL approach agents use econometric forecasting rules to form expectations of the key variables y^e_{t+s} and π^e_{t+1}, as well A^e_{t+s} and v^e_{t+s}, that are required for decision-making, using a forecasting model that depends on the observed exogenous shocks and any observed predetermined aggregate state variables. The forecasting model is often assumed, in linear set-ups, to take a form that nests the RE equilibrium (REE), but with coefficients that are unknown and so must be estimated using observed data. Under AL the required parameter coefficients are updated over time in line with data generated by the evolving equilibrium path. In linear models the forecast model can usually be chosen to nest the REE of interest, so that AL has the potential to converge to RE over time.

Apart from providing a plausible form of bounded rationality in macroeconomic economic models, AL is particularly of interest for two reasons. First, it provides a natural stability condition for an REE. Provided the REE satisfies the E-stability condition for a suitable forecasting specification – known as the "perceived law of motion" (PLM) – the REE is locally stable under AL. If the REE is expectationally stable ("E-stable") the PLM will (locally) converge over time to RE (or to a neighborhood of RE when agents use "constant gain" RLS). E-stability conditions are typically straightforward to derive.[2]

In many stochastic nonlinear models a closed-form RE solution may be unavailable, but a suitable linear forecasting rule may converge to a solution – a "restricted perceptions equilibrium" (RPE) – that approximates REE. If the REE is E-stable but agents use linear forecasting rules, there will typically be convergence to an RPE corresponding to the REE. Stability under learning is particularly of interest in models with multiple REE, as in the current model. The learning dynamics themselves are often of interest, both in stationary environments and most notably in the context of policy changes.

3 Nonlinear NK model with Adaptive Learning and Cognitive Discounting

To incorporate CD into our AL framework we start with the simplest case, without stochastic shocks, in which agents use "steady-state learning" that recursively estimates the means of output and inflation according to

$$y^e_t = y^e_{t-1} + \omega\left(y_{t-1} - y^e_{t-1}\right) \text{ and } \pi^e_t = \pi^e_{t-1} + \omega\left(\pi_{t-1} - \pi^e_{t-1}\right),$$

where $0 < \omega < 1$ is the learning "gain" parameter. Usually we focus on small fixed gains $\omega > 0$.[3]

[2] For a general discussion and analysis of adaptive learning, expectational stability and E-stability dynamics, see Evans and Honkapohja (2001). Discussion of more recent literature and applications can be found in Evans and McGough (2020)

[3] Decreasing gains in which ω_t is proportional to t^{-1} are important for some theoretical results.

With AL we include CD by discounting AL expected future output and inflation at rate $0 \leq m_y \leq 1$ and discounting of AL expected future real interest rates at rate $0 \leq m_r \leq 1$. This is most conveniently formulated mathematically in a multiplicative set-up. Let $x^* \in \{\pi^*, y^*\}$ be the target towards which future forecasts are discounted at rate m_x, where $x \in \{y, \pi\}$. For the steady-state learning case, the forecast $n-$step ahead forecast x^e is discounted geometrically at rate m_x toward the target:

$$x_n^e = x^* \left(\frac{x^e}{x^*} \right)^{m_x^n}.$$

Note that for $m_x = 1$ future forecasts $x_n^e = x^e$ are not discounted and for the other polar case $m_x = 0$, forecasts are fully discounted to the target x^*. This multiplicative framework for CD is particularly useful computationally, because it allows for results on Mahler series that can be exploited in the formulation of the temporary equilibrium equations.

In the Appendix it is shown that in the non-stochastic set-up with $A_t = A$ and $\nu_t = \nu$, the "temporary equilibrium" (TE) equations, giving y and π as functions of y^e and π^e, are provided by the following "IS" and "AS" equations.

$$\begin{aligned}
\text{IS} \quad : \quad & y = (1-\xi)g + (\beta^{-1} - 1) \left(\frac{r^* \pi^e}{R(\pi^e, y^e)} \right)^{\frac{m_r}{1-m_r}} \\
& \times \left\{ \begin{array}{c} \left(y^* \sum\limits_{n \geq 1} \beta^n \left(\left(\frac{R(\pi^e, y^e)}{r^* \pi^e} \right)^{\frac{m_r}{m_r - 1}} \right)^{m_r^n} \left(\frac{y^e}{y^*} \right)^{m_y^n} \right) \\ -(1-\xi) g \sum\limits_{n \geq 1} \beta^n \left(\left(\frac{R(\pi^e, y^e)}{r^* \pi^e} \right)^{\frac{m_r}{m_r - 1}} \right)^{m_r^n} \end{array} \right\} \\
\text{AS:} \quad & Q(\pi) = G(T_y(\pi^e, y^e)) + \left(\frac{\nu}{\alpha} \right) \left(\frac{y^*}{A} \right)^{\frac{1+\varepsilon}{\alpha}} \sum_{n \geq 1} \beta^n \left(\left(\frac{y^e}{y^*} \right)^{\frac{1+\varepsilon}{\alpha}} \right)^{m_y^n} \\
& + (1-\nu) \sum_{n \geq 1} \beta^n \left(\frac{y^* \left(\frac{y^e}{y^*} \right)^{m_y^n}}{y^* \left(\frac{y^e}{y^*} \right)^{m_y^n} - (1-\xi) g} \right)
\end{aligned}$$

Note that, given the CD parameters, in the nonstochastic case under consideration, y is a function of (π^e, y^e). In the AS equation, $T_y(\pi^e, y^e)$ is defined implicitly as the value of y that solves the IS equation for given (π^e, y^e). Likewise we can define $T_\pi(\pi^e, y^e)$ as the value of π that solves the AS equation.

The IS and AS curves define the temporary equilibrium map from expectations (π^e, y^e) to realized (π, y). Once calibrated, approximations of this map can be used to provide phase diagrams and stochastic simulations. Naive truncation – replacing the series with

partial sums – introduces numerical instability. Under appropriate transformation, results on Mahler series can be exploited to obtain code that delivers fast, stable computations: see Appendix A for details.

4 Steady States and E-stability Dynamics

CD can have substantial quantitative and qualitative effects on the phase diagram for (π^e, y^e). We will focus on the *Baseline Calibration* given in Table 1. Most parameter values are in line with the IER paper, but we now include CD parameters m_r and m_y for the reasons discussed above: long-range forecasts based on estimated econometric models are likely to be less reliable than short-range forecasts because of nonstationarities in the data including in particular unpredictable structural shifts. See, for example, Chapter 21 of Hendry and Nielsen (2007). In this context, the arguments of Gabaix (2020) for discounting distant forecasts to a suitable steady state are particularly strong, and in the context of our model, in which policymakers aim to achieve full-employment output together with a low, positive rate of inflation, the targeted steady-state levels of y and π is the natural choice.

Because our nonlinear NK model has *three* steady states, a natural question to ask is whether CD can alter the positions of the two non-target steady states. In particular, is it possible for there to be a low output steady state that is stable under adaptive learning? To sharpen this question, we investigate whether, with suitable CD parameters there will exist a "stagnation" steady state in which inflation and output gaps are in line with the trough of the US Great Depression in the US. Specifically we look for CD parameters that would match a steady state with a quarterly inflation factor of about 0.965 (corresponding to a deflation rate of 10% per year) and an output gap factor of about 0.73, corresponding to the 1933 trough, which was 74% of its the 1929 GDP peak.

Table 1 gives the parameters for our "*Baseline Calibration*". Most of the parameter values are in line with those from the IER paper, but we now incorporate CD parameters m_y, m_r chosen to deliver a steady state with output and inflation in line with the US Great Depression.

4.1 Phase diagram for Baseline Calibration and the Great Depression

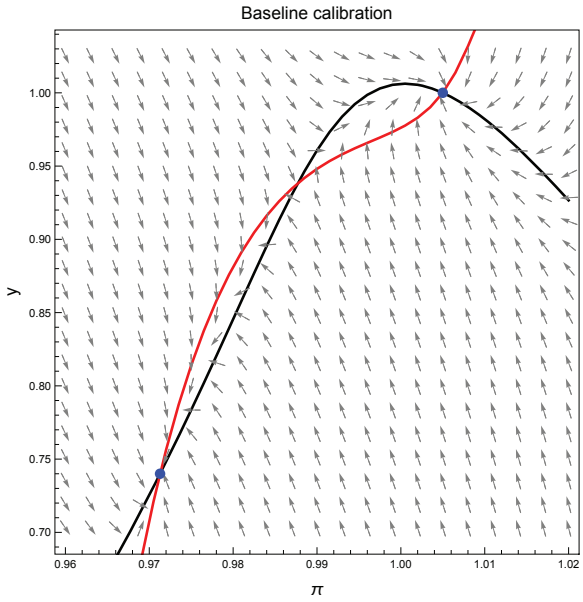
Figure 1 gives the phase diagram for expectation variables π^e, y^e induced by AL and CD with the baseline calibration. For clarity, only part of its DOA (domain of attraction) is shown. There are three steady states. The "targeted" steady state (top right) corresponds to $y = 1$ and $\pi = 1.005$ (i.e. 2% annual inflation target). It can seen that it is locally E-stable. The black null-cline corresponds to $dy^e/dt = 0$ and the red null-cline to $d\pi^e/dt = 0$.

Table 1: Model Parameters and Policy Variables

Parameter	Value	Description
Model Parameters		
A	1.113	Productivity factor
ρ_A	0.95	Productivity autocorrelation (logs)
β	0.99	Discount factor
α	0.7	Capital share
ν	13.5	Elasticity of substitution between varieties
ρ_ν	0.95	Elasticity of substitution autocorrelation (logs)
ε	1.0	Frisch elasticity
ξ	0.4	Private/public substitution parameter
Cognitive Discounting Parameters		
m_r	0.94	Interest memory retention
m_y	0.8265	Output memory retention
Adjustment Cost Parameters		
ϕ	75.0	Adjustment cost factor
ψ	20.0	Elasticity of adjustment kinda
ζ	0.475	Smoothing parameter
π_{LL}	0.8395	Hard lower bound on inflation
π_L	0.95	Soft lower bound on inflation
π_H	1.015	Soft upper bound on inflation
π_{HH}	1.1	Hard upper bound on inflation
Policy Parameters		
π^*	1.005	Inflation target
R^*	π^*/β	Steady-state interest rate
B	$1.5/R^*$	Inflation elasticity
ϕ_y	8.2	Output elasticity
$\bar{g}$	0.2	Steady state Government spending
$\bar{r}$	$1/\beta$	Steady state real interest rate

The well-known "unintended" middle steady state, induced by the ZLB, is approximately at $\pi^e = 0.988$, $y^e = 0.94$, but is *not* E-stable and hence not stable under learning. More

Figure 1: Phase space



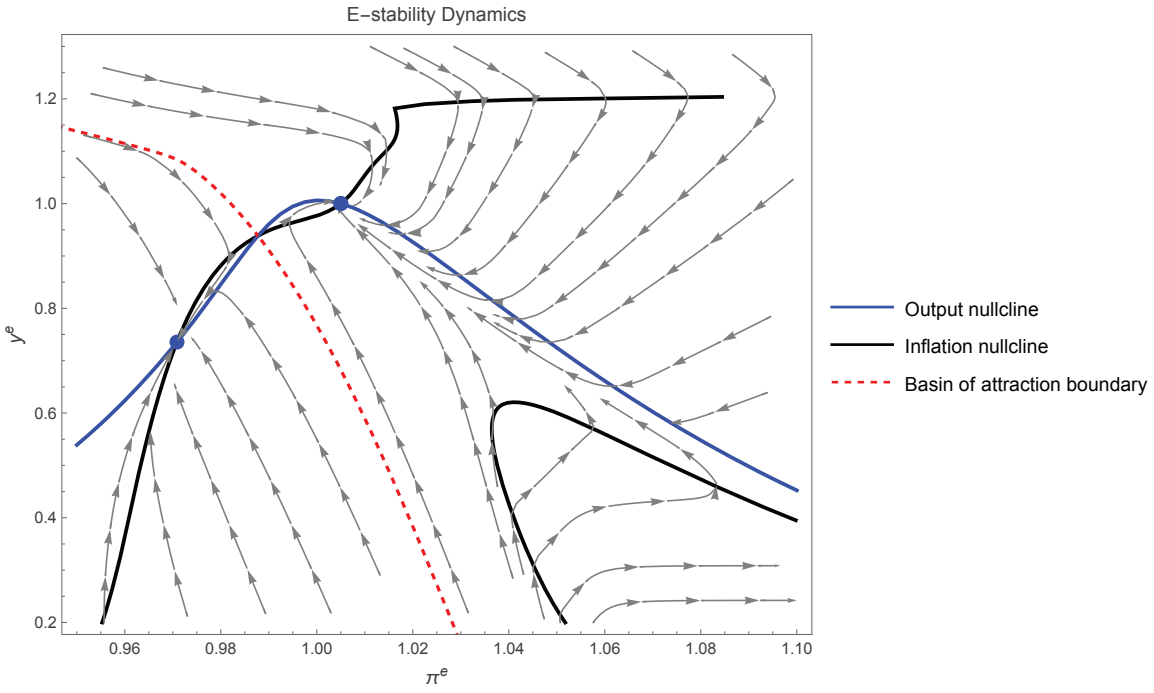
precisely, it is a saddlepoint that lies on the boundary between the two basins of attraction. The phase diagram in Figure 1 shows two locally stable steady states: the targeted steady state and the stagnation steady state, each of which has a sizable basin of attraction. Figure 2 below shows the boundary between the two basins of attraction as a red dashed curve; formally the boundary between two DOAs is on a saddle path that converges to the middle steady state.

The low "stagnation" steady state at ($\pi = 0.971$ and $y = 0.740$) matches the trough of the 1930s US Great Depression; in 1933 y was 74% of its 1929 peak value, and the maximum annual deflation factor (in 1934) was 0.90 (a 10% deflation rate, corresponding to $\pi = 0.975$ quarterly). In can be seen from Figure 1 that the low stagnation steady state *is E-stable*, and hence stable under adaptive learning. The CD parameters $m_r = 0.94$ and $m_y = 0.8265$ chosen are roughly in line with numerical values suggested by Gabaix of 0.8 and 0.85.

What lies behind the existence and local stability of the stagnation steady state? Intuitively, at the stagnation steady state, under pure AL without CD, the temporary equilibrium levels of output and inflation would both be lower than the stagnation steady state values. However because of CD, long-horizon forecasts of y and π are discounted toward target values, offsetting the more extreme short-term pessimism that would otherwise lead to a persistent decline in output and inflation. The stagnation steady state, and its local domain of attraction, reflects this balance.

Figure 2 shows the global situation for the whole range of output and a wide range of inflation. In some regions the expectation dynamics are complex, but globally there are

Figure 2: Global flow diagram



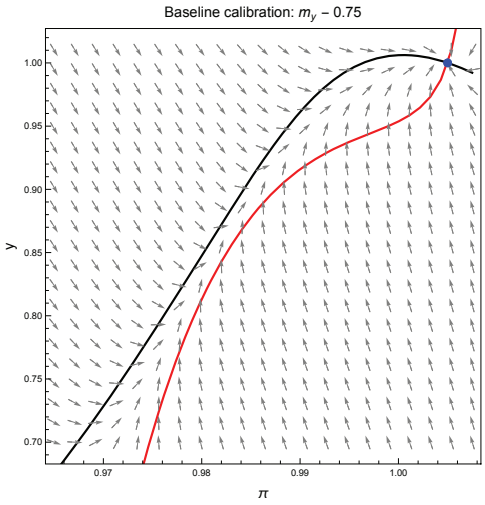
just the three steady states: the locally E-stable target steady state, the E-unstable "middle" steady state, and the locally E-stable Great Depression steady state. Also shown in Figure 2 is the boundary separating the basins of attraction of the two stable steady states.

4.2 Alternative CD calibrations

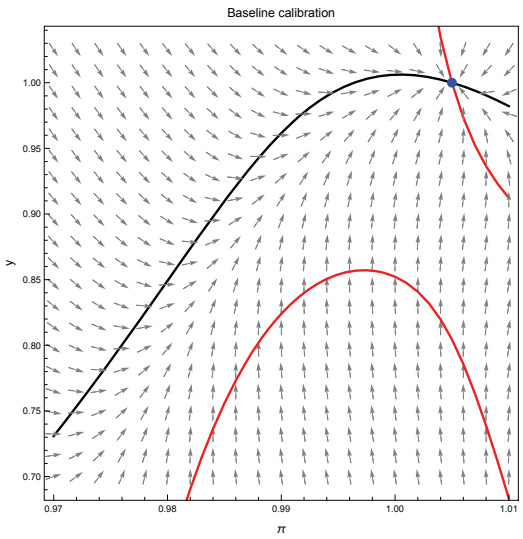
The qualitative features of the phase diagram, including the number of steady states, depend on the CD parameter values, in particular on m_y. Changing m_y to $m_y = 0.75$, gives the phase diagram shown in Figure 3a and for $m_y = 0.6$ the phase diagram shown in Figure 3b. In both of these cases the model has a unique, globally E-stable steady state at the policy target. In Figure 3a convergence to the target steady state is locally cyclical, while in Figure 3b the learning dynamics becomes locally monotonic. These results are not surprising: when CD is very strong (low m_y) longer horizon expectations y^e and π^e are discounted towards a path of rapid convergence to the target steady state, leading to a unique E-stable steady state at the policy target. Essentially, the intertemporal aspects (y^e, π^e) are rendered more monotonic by CD of the model and multiple steady states can no longer exist.[4]

[4]Note that the red isocline has a single connected component in Figure 3(a) and appears to have two connected components in Figure 3(b). This bifurcation is discussed in Appendix E.

Figure 3: Other steady-state configurations



(a) $m_y = 0.75$



(b) $m_y = 0.6$

4.3 Further discussion

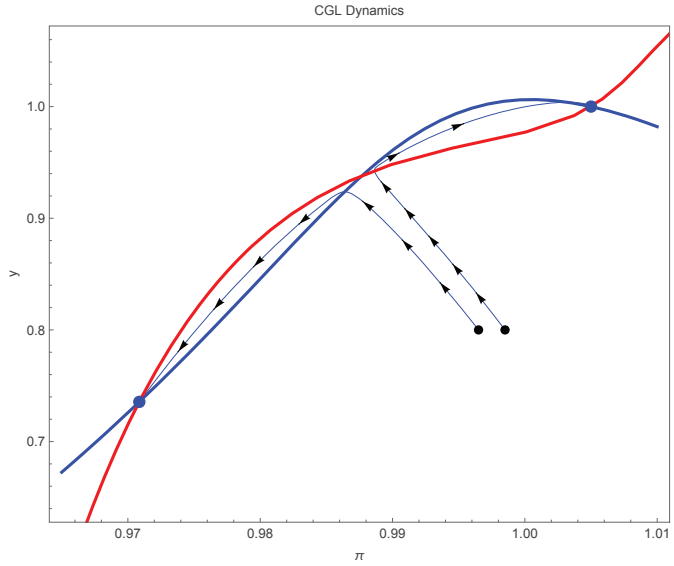
Returning to the baseline Great Depression calibration, it is can be seen from Figure 2, giving the global expectation dynamics, that for initial expectations near the DOA boundary, there is strong sensitivity to initial conditions. Specifically, for initial (π^e, y^e) ; just North East of the boundary, there will be convergence to the target steady state, while for initial (π^e, y^e) just South West of the boundary, there will be convergence to the Depression steady state. Figure 4 shows the contrasting time paths for these two nearby starting points.

We note that the paths shown in Figure 4 capture the dynamics under AL with CD when there are no stochastic shocks in the economy. When exogenous stochastic shocks are included the resulting paths will be stochastic, reflecting the impact of these shocks. The paths shown can still be interpreted as the mean paths, but the stochasticity implies that for either of the two nearby starting points, there can be eventual convergence to either of the two E-stable steady states, depending on the sequence of shocks.

5 AL and CD in a Stochastic Economy: net IRFs

Using our baseline calibration, and now including stationary $AR(1)$ shocks to A_t and υ_t, we can compute, starting, e.g. at the targeted steady state, net impulse response functions (IRFs) for shocks to A_t or ν_t. To do this computationally we simulate multiple independent

Figure 4: Sensitivity of dynamics to initial conditions



stochastic time paths following a shock to A_t or ν_t and compute the average impact on over time on π_{t+s} and y_{t+s} of an initial departure of A_t or υ_t from its mean. We set the AR(1) parameters for $\log A_t$ and $\log \pi_t$ at 0.95. The first figure shows the net IRFs for $\log y_{t+s}$ and $\log \pi_{t+s}$ of a 1% increase in productivity at time zero. Black curves denote realized y and π, and red curves denote y^e and π^e. The (net) IRF for y reflects the short-run positive effect of productivity on output, and a negative effect on inflation arising from the reduced shadow marginal cost, with effects converging asymptotically to zero over time as the economy returns to the steady state.

Because our model is purely forward-looking, y_t and π_t depend only on current exogenous shocks A_t, ν_t, g_t and on forecasts y^e_{t+s} and π^e_{t+s},for $s = 1,2,3,\ldots$. We will limit attention here to observable shocks including in particular A_t, ν_t, that follow stationary AR(1) processes.

$$A_t = \bar{A}^{1-\rho_A} A_{t-1}^{\rho_A} \varepsilon_t^A \text{ and } \upsilon_t = \bar{\upsilon}^{1-\rho_\upsilon} \upsilon_t^{\rho_\upsilon} \varepsilon_t^\upsilon .$$

where ε_t^A, ε_t^υ are *iid*, strictly positive and unit mean. Here we assume agents use point expectations of the exogenous shocks, so

$$A^e_{t+n} = \bar{A}\left(\frac{A_t}{\bar{A}}\right)^{\rho_A^n} \text{ and } \upsilon^e_{t+n} = \bar{\upsilon}\left(\frac{\upsilon_t}{\bar{\upsilon}}\right)^{\rho_\upsilon^n}.$$

Note that these shocks only enter the marginal cost term $G(y_t, A_t, \upsilon_t) = \frac{\upsilon_t}{\alpha}\left(\frac{y_t}{A_t}\right)^{\frac{1+\varepsilon}{\alpha}} - \frac{y_t(\upsilon_t - 1)}{y_t - (1-\xi)g}$.

Assume a fixed level of government spending $g_t = \bar{g}$, and that, as in a stochastic steady-state REE, expected output and expected inflation are functions of A_t, ν_t. For expectations

we have

$$G(y^e_{t+n},A^e_{t+n},\upsilon^e_{t+n}) = \left(\frac{\upsilon}{\alpha}\right)\left(\frac{y^*}{\bar{A}}\right)^{\frac{1+\varepsilon}{\alpha}}\left(\frac{\upsilon_t}{\bar{\upsilon}}\right)^{\rho^n_\upsilon}\left(\left(\frac{A_t}{\bar{A}}\right)^{-\frac{1+\varepsilon}{\alpha}}\right)^{\rho^n_A}\left(\left(\frac{y^e_{t-1}}{y^*}\right)^{\frac{1+\varepsilon}{\alpha}}\right)^{m^n_y}$$
$$-\frac{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}\bar{\upsilon}\left(\frac{\upsilon_t}{\bar{\upsilon}}\right)^{\rho^n_\upsilon}}{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}-(1-\xi)g}+\frac{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}}{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}-(1-\xi)g}$$

and the Phillips curve is

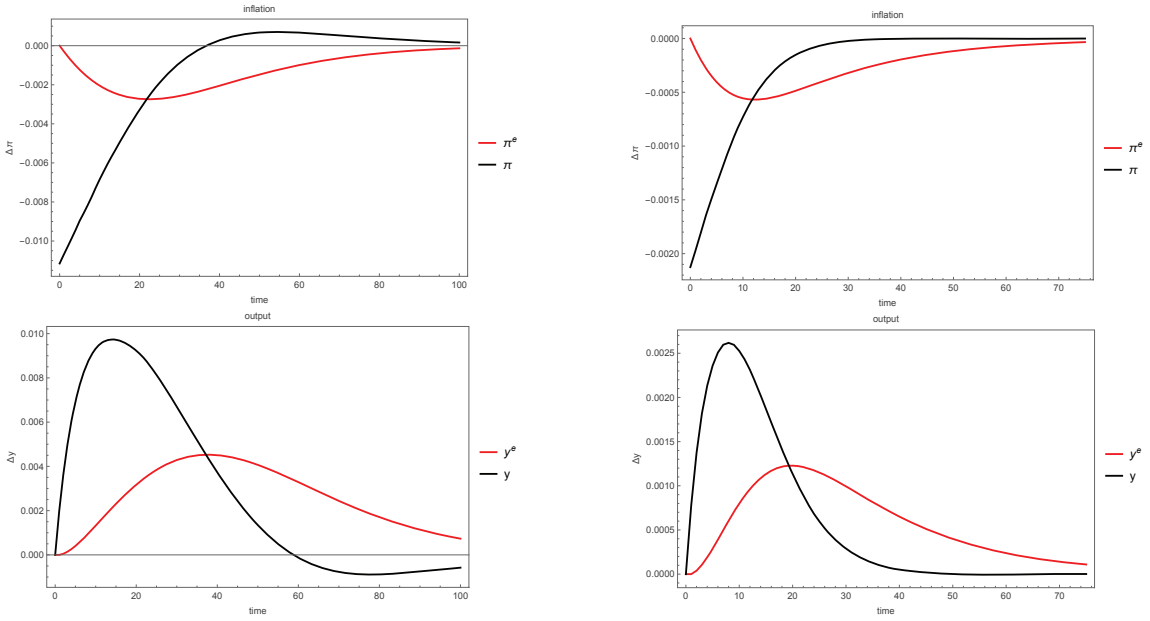
$$\begin{aligned}\phi'(\pi)\pi &= G(y_t)+\sum_{n\geq 1}\beta^n G(y^e_{t+n},A^e_{t+n},\upsilon^e_{t+n})\\ &= G(y_t)+\left(\frac{\bar{\upsilon}}{\alpha\bar{A}^{\frac{1+\varepsilon}{\alpha}}}\right)\sum_{n\geq 1}\beta^n\left(\frac{\upsilon_t}{\bar{\upsilon}}\right)^{\rho^n_\upsilon}\left(\left(\frac{A_t}{\bar{A}}\right)^{-\frac{1+\varepsilon}{\alpha}}\right)^{\rho^n_A}\left(\left(\frac{y^e_{t-1}}{y^*}\right)^{\frac{1+\varepsilon}{\alpha}}\right)^{m^n_y}\\ &\quad -\bar{\upsilon}\sum_{n\geq 1}\beta^n\left(\frac{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}\bar{\upsilon}\left(\frac{\upsilon_t}{\bar{\upsilon}}\right)^{\rho^n_\upsilon}}{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}-(1-\xi)g}\right)+\sum_{n\geq 1}\beta^n\left(\frac{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}}{y^*\left(\frac{y^e_{t-1}}{y^*}\right)^{m^n_y}-(1-\xi)g}\right).\end{aligned}$$

The dynamics of the stochastic economy can be illustrated by numerically computing IRF that arise from different random shocks. Consider first a 1% shock to productivity in period zero. Productivity is calibrated with log productivity having serial correlation 0.95 and gain 0.01. Two innovations to productivity are obtained by assuming that the initial shock is zero or one percent. These sequences are incorporated to the model to generate two time series which are then compared by computing their difference. The differenced series give the response of the model from initial steady state and conditional on the drawn series of innovations. The final step is eliminate conditional dependence by running 10000 simulations and averaging these outcomes to eliminate conditional dependence. Figure 5(a) gives results for inflation and expected inflation as well as output and expected output.

The two panels show the net IRFs for inflation and output from a 1% increase in productivity at time zero. Black curves denote realized y and π, and red curves denote y^e and π^e. The (net) IRF for y reacts the short-run positively of productivity on output, and a negative effect on inflation arising from the reduced shadow marginal cost, with effects converging asymptotically to zero over time as the economy returns to the steady state.

Figure 5(b) shows the effects of elasticity of substitution shocks using net IRFs for inflation and output of a time zero 1% increase in the elasticity of substitution across goods. This corresponds (approximately) to a one percentage point decrease in the mark-up. This results in the short term with a decrease in inflation and, because of the monetary policy response, with decreased interest rates and higher output. Again, these effects disappear asymptotically as the economy returns to the target steady state.

Figure 5: IRFS



(a) Productivity

(b) elasticity of substitution

5.1 Escape paths (preliminary and incomplete)

One of our more striking findings is the possibility of a suitable sequence of random shocks leading to an **escape path** from the targeted steady state to the stagnation steady state. The likelihood of this is higher when the AR(1) parameters $0 < \rho_A, \rho_v < 1$ are close to one.

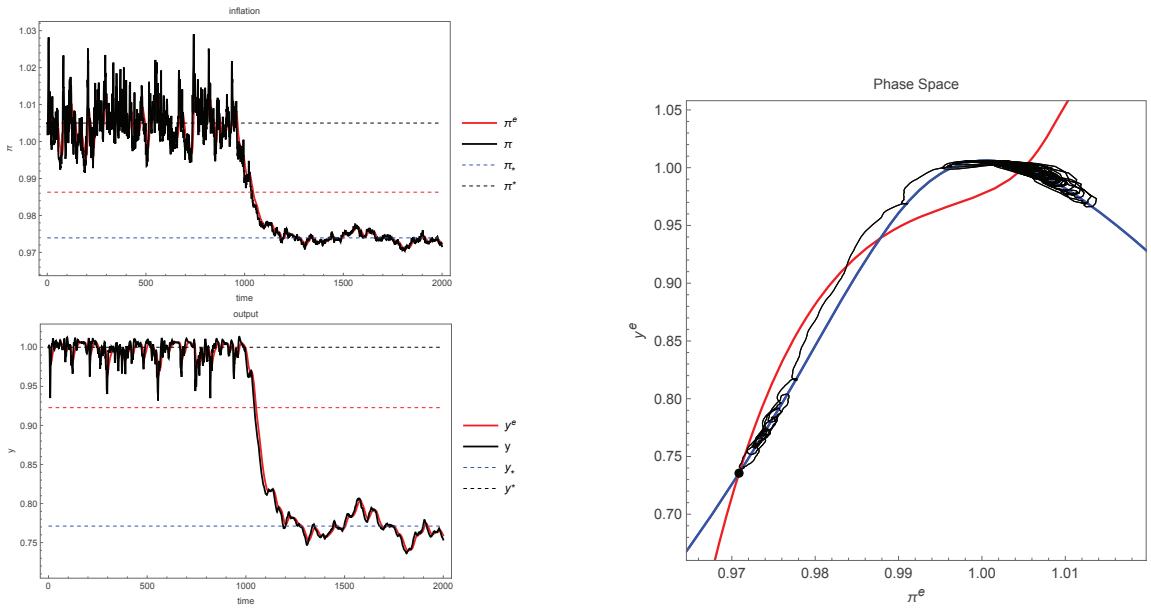
The possibility of escapes is illustrated in the following Figure. Figure 6(a) shows an example of an endogenous "escape" of the economy, from the target steady state to the stagnation steady state, as a result of an unfortunate sequence of exogenous shocks. Figure 6(b) shows the same time path plotted in (π, y) space.

However, it is rare for the economy to "escape" endogenously from the Depression steady state to the targeted steady state as a result of a fortunate sequence of exogenous shocks. This raises the question of whether *policy* can move the economy from the Depression steady state to the target steady state.

5.2 A Digression on Gains (preliminary and incomplete)

Under recursive AL a key parameter is the "gain" parameter, which governs the weight placed on the most recent data point when updating the parameters of the forecast rule. In 1-step ahead models, like the textbook 2-equation model above, in which y, π depend

Figure 6: Escape path from target steady state to stagnation



(a) Time series

(b) Phase space

on next period y^e, π^e, typical (constant) gains used have been in the range 0.015 to 0.05. For example, Branch and Evans (2006) found the best fit both for quarterly output growth and inflation forecasts, as well as for matching professional forecasters was $(0.015, 0.05)$. Milani (2007), in an estimated quarterly NK model with AL obtained a posterior estimate for the constant gain of 0.0183. Orphanides and Williams (2002) and Orphanides and Williams (2005) also used quarterly gains in the interval $(0.015, 0.03)$.

In contrast the gains used in long-horizon AL models are usually *much* smaller. For example, in the paper Eusepi and Preston (2011) the gain used was 0.002, an order of magnitude smaller. Our framework casts some light on this issue. In our long-horizon model, without CD, and starting near the targeted steady state, with $y^e = 1$ but $\pi^e = 1.0075$ (3% annual), we find that gains above 0.03 start to exhibit instability. As we increase gains from 0.03 to 0.05 or higher, the economy exhibits large fluctuations, large damped 3-cycles and divergence. Similar issues arise when the economy is subject to a one period 5% increase in government spending g.

This suggests that in long-horizon models without CD, the present-value sums of expected future output and inflation are implausibly sensitive to small changes in expectations. However, this "oversensitivity" goes away for CD incorporated with, for example, $m_y = m_r = 0.85$.

6 Fiscal Policy

Are there policies that can push the economy out of a Depression steady state (or that would avoid Depression when expectations turn pessimistic)? One natural candidate is aggressive fiscal policy. In EHM, we found that, in many cases, an announced expansionary fiscal policy could avoid stagnation provided expectations were not yet too pessimistic. In the current paper we set a more ambitious goal. Are there policies which, if the economy is in already in a Depression steady state, will successfully push the economy out of the DOA (domain of attraction) of the depression steady state, and leaded to eventual convergence to the target steady state.

A natural policy candidate is an announced and implemented fiscal policy that increases government spending by a substantial amount for an extended period of time. This can be studied using the approach introduced and studied in Evans, Honkapohja, and Mitra (2009): it is assumed that agents understand the policy, treat it as credible, incorporate the direct effect of the announced policy in their long-horizon decision rules, and use AL to revise their forecasts y^e_{t+s} and π^e_{t+s} over time using the evolving paths of y_t, π_t. See Appendix B for technical discussion.

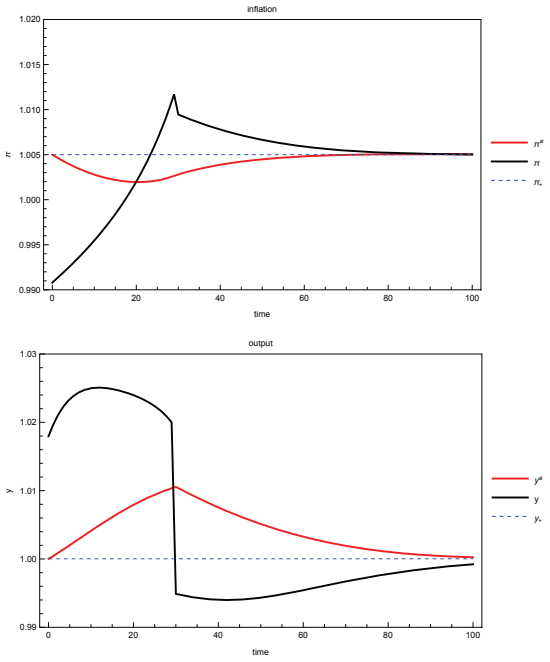
In our current setting, we are particularly interested in whether, starting from the more extreme case of an economy in the Depression steady case, a credibly announced and implemented policy that increases g to $g + \Delta g$ for T periods, (after which government spending reverts to g, can push the economy into the DOA of the target steady state. It is assumed that agents understand this policy and incorporate it into their decision-making. Before studying this, it is useful to examine the impact of such a policy starting from the target steady state.

6.1 Anticipated expansionary fiscal policy starting from target steady state.

Our general framework, with AL and CD, can be readily extended to incorporate the anticipated increase in government spending into agents' decision-making, so that dynamics under anticipated policy can be determined as usual using our computationally efficient temporary equilibrium maps. To get some intuition for the mechanisms in play, we start by looking at the net IRFs for an announced an implemented increase in g, starting from the target steady state.

Figure 9 shows the impact on π and y of a 20% increase in g for ten periods,starting from the target steady state (beliefs in red & realized values in black). Although output initially increases substantially during the fiscal stimulus, inflation becomes negative for the first six periods. Intuition for short-term deflation despite increased output is along lines familiar

Figure 7: Anticipated fiscal stimuluous



from RBC models: increased taxes used to finance the period of higher government spending reduces the present value of disposable income, which leads agents to reduce leisure as well as consumption. The corresponding increased supply of labor results in lower shadow marginal costs and hence lower prices.

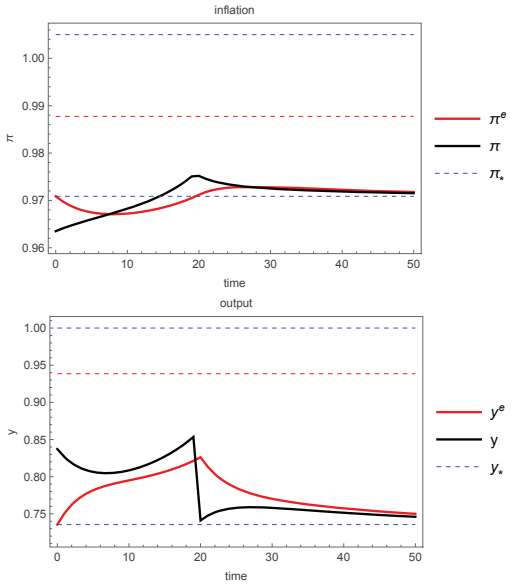
6.2 Anticipated expansionary fiscal policy in a depression

To see whether an anticipated extended large increase in government spending g can push the economy from the Depression steady state to the target steady case, we consider an anticipated increase in g that doubles it size, from $g = 0.2$ to $g = 0.4$ for 20 periods: see Figure 8.

The large anticipated expansionary fiscal policy does not succeed in shifting the economy from stagnation to the targeted steady state. The explanation for its lack of success is the disinflationary aspect of anticipated fiscal expansions, noted above. The policy to shift the economy from a Great Depression steady state to the target steady state fails because it does not increase inflation as well as output.

This can be understood by returning to Figure 2. To escape the DOA of the Depression steady state, policy needs to shift (π^e, y^e) to the north-east of the dashed red line. This would be difficult to achieve without an increase in π^e, and, under adaptive learning, this

Figure 8: Effects of increase in g starting from Depression steadystate



would require a sustained increase in actual inflation π. However, the labor supply effect, familiar from RBC models, *reduces* inflation. Consequently the policy is unsuccessful.

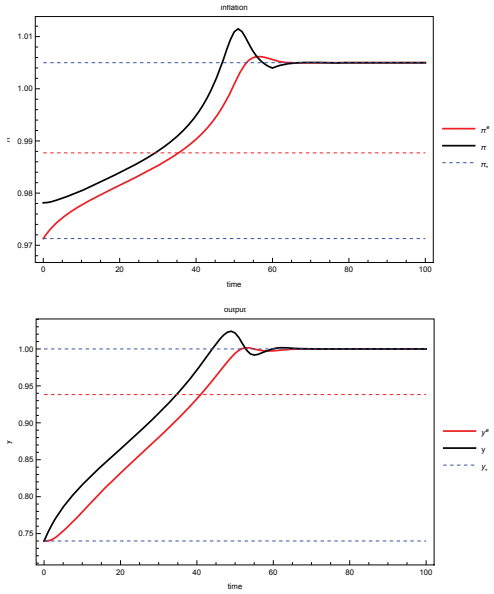
7 Policies to increase prices and wages

During the New Deal, in the US, a variety of policies were put in place to ameliorate the effects of the Great Depression and to push the economy toward recovery. While ultimately the later huge increase in g associated with WWII led to a full economic recovery, there were other salient policies that may have been important for the recovery. In particular, in addition to emergency measures to support the unemployed and avoid foreclosures, and financial measures to prevent runs on banks, policies were introduced to support wages and prices.

In an effort to combat deflation, a range of policies were introduced to induce a rise in the price level, including:
– policies to support agricultural prices
– industry codes to eliminate unfair competitive practices; establishment of minimum wages and maximum hours; support for labor's right to collective bargaining
– gold policy: dollar devaluation reducing the gold content of the $ to 59% of its former value
– defense of organized labor, increased unionization and support for collective bargaining & creation of NLRB

The importance of these and other factors is examined in Eggertsson (2008) and Eggertsson

Figure 9: Effects of increase in mark-up starting from Depression steady state



(2012). Eggertsson (2008) stresses these policy changes within the context of an overall "policy regime change" that eliminated previous "policy dogmas," engineering a "regime change" that shifted expectations from contractionary to expansionary. Eggertsson (2012) takes a different tack, arguing that policies to increase nominal prices and wages had a more direct effect by reducing deflation and consequently ex-ante real interest rates.

Our model formalizes and captures these channels both through direct impacts and, over time, via the AL channel, as discussed in Section 1.4 of the Introduction. A simple and natural way to model these policies is to view them as facilitating a sustained reduction in ν, i.e. a credibly announced policy of a sustained increase in the mark-up.

7.1 Anticipated policies to increase mark-ups in a Depression.

Assume that the economy is in the Depression steady state. Policymakers then announce and implement policies that decrease ν by 20%, so that there is an increase in the mark-up by approximately 20%, for 60 periods, after which ν returns to its original level. As with the fiscal policy experiment, it is assumed that agents understand this policy and incorporate it into their decision-making.

Both π and y increase steadily over time for about 50 periods, both somewhat overshooting the targeted steady state values, and then following a damped cyclical path converging to the target steady state. The increase in mark-ups increases inflation, i.e. reduces deflation rates, and, through adaptive learning, expected deflation rates. With nominal interest rates

close to the zero lower bound, reductions in expected deflation reduced the expected real interest rate. This leads to higher demand and higher output, which via adaptive learning leads to higher expected output. This reinforcing feedback loop pushes the economy into the DOA of the target steady state, with eventual convergence to the target steady state.

This is a powerful and striking result. Actual and expected deflation, at the stagnation steady state, is a major impediment to recovery from Great Depression levels of output. Policy induced increased mark-ups that decrease deflation at the Depression steady state, can increase π^e and y^e enough to put expectations into the domain of attraction of the target steady state so that after the policy change ends there will be convergence to the target steady state.

8 Expansionary Fiscal Policy Redux

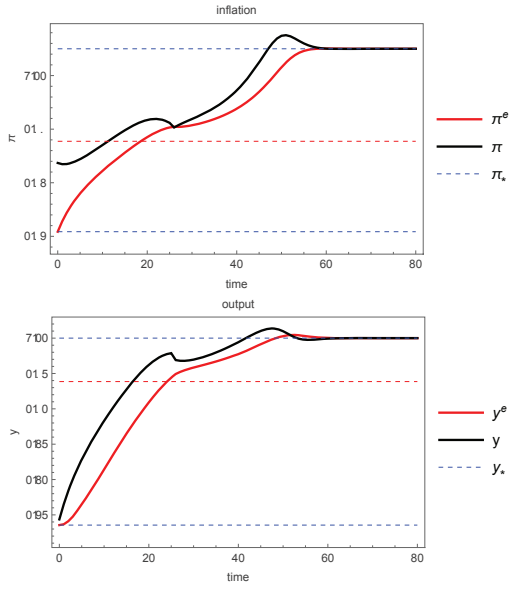
In Section 6 we found, starting from the Depression steady state, that an anticipated aggressive fiscal expansion, from $g = 0.2$ to $g' = 0.4$, failed to return the economy to the target steady state. The underlying reason for this was that while the increased government spending increased demand and initially led to higher output, the higher anticipated taxes to fund the $\uparrow g$ led to a negative income effect, which induced agents to reduce consumption of all goods, including leisure. This increased labor supply, leading initially to reduced wages, and hence lower marginal costs and lower prices, with little net impact on deflation over the period of higher g, leading to a return to the Depression steady state.

We now consider whether results can change significantly for alternative choices in the composition of the fiscal stimulus. Specifically, we consider the impact of an announced increase in government spending g, if the composition of the announced $\uparrow g$ takes the form of government expenditure on goods that are close substitutes to private consumption, i.e. we look at an announced $\uparrow g$ for which $\xi' > \xi$ for government spending during this period.

Because the income effect is $-\frac{y(\nu-1)}{y-(1-\xi')g}$, when ξ' is close to one the size of the negative income effect is greatly reduced. Figure 12 looks at the effect, starting in the Depression steady state, of an announced doubling of government spending (from 0.2 to 0.4) for 26 periods (quarters) in which, during this period, the additional govt spending has $\xi' = 0.94$, i.e. the additional government spending is on goods that are close substitutes for private consumption.

Figure 12 shows that the large expansionary fiscal policy with this composition is now successful in our model. This type of government spending may not have played a big role for the US in the New Deal, except for, say, the $\uparrow$ availability of electricity and flood control in the South made possible by the TVA. However, it is easy to think of contemporary examples, such as public spending on or financing of child care, public schools, SNAP (food stamps), health care, and national parks (for recreation).

Figure 10: Anticipated increase in g, starting from Depression steady state, for spending on close substitutes to private consumption



9 Anticipated bottlenecks: preliminary and incomplete

We model an anticipated, temporary supply shock as an "announcement" in period t_0 of a step decrease in productivity from A to $\hat{A}$ for periods $t_1 - t_2$, after which productivity returns to its calibrated value.

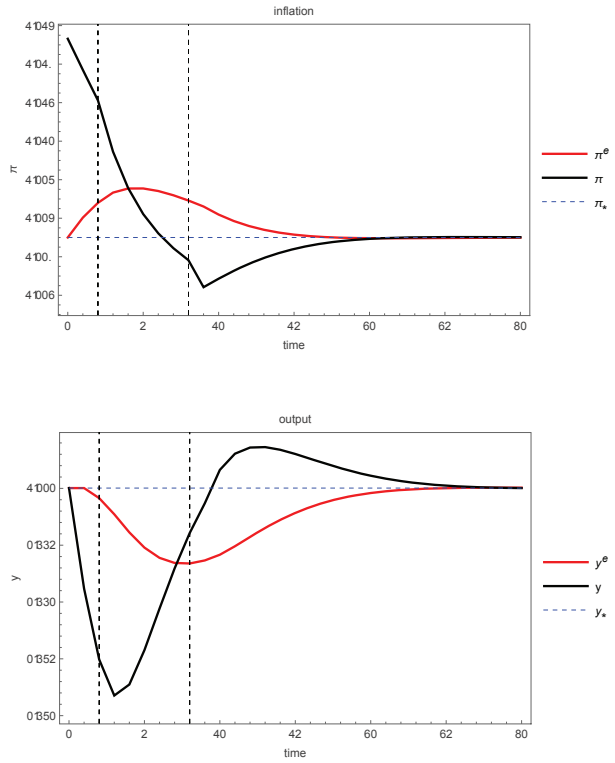
A fall in productivity raises marginal costs thus putting upward pressure on inflation; higher inflation leads to higher expected inflation, thus raising real interest rates and lowering output. Because firms are forward-looking and want to smooth price changes, even an anticipated fall in productivity raises inflation, thus initiating the attendant fall in output (with one period lag since output depends only on expectations).

The figure below provides the following experiment: in period $t_0 = 0$ it is announced that in period $t_1 = 2$ productivity will fall by 1.0%, and will remain low for six periods $(t_2 = 8)$. The needed analysis is found in Appendix C.

10 Extensions

The analytical framework of this paper can incorporate a wide range of exogenous shocks. Major events in the US and other major market economies include the oil price shocks that led first to high inflation and then to stagflation in the late 1970s and early 1980s.

Figure 11: Anticipated fall in A



Financial shocks precipitated the Great Recession of 2008-9, which was followed by an unusually slow recovery. Labor supply and bottleneck shocks during the Covid Pandemic Recession (March-April 2020 in the US), together with expansionary fiscal and monetary policies, was followed by over two years of elevated inflation. Beginning in early 2025, the US announced large increases in tariffs against many countries, which, however, been frequently followed by changes in the announced rates and considerable uncertainty about tariff rates going forward.

An advantage of the multiplicative structure of our model is that it facilitates inclusion of a wide range of additional exogenous shocks. In work in progress we are examining the implications of including (i) Finite-state stochastic discount rates, to capture exogenous changes in consumer/firm confidence, (ii) financial shocks, modeled as a wedge between borrowing and lending rates, (iii) national sales taxes, acting as a stand-in for tariffs, and (iv) reduced labor supply, due to Covid shocks, combined with reduced productivity arising from related sectoral bottlenecks. Appropriate policy in the presence of one or of a combination of several large exogenous shocks of these types, would clearly be of interest.

Variations of our current framework can also be examined. In line with much of the literature, we make the assumption that the interest-rate rule used by policymakers is a known

function of (π^e, y^e). One justification of this is that it captures the importance of transparent monetary policy rules, a point stressed by Eusepi and Preston (2010). However, a natural alternative assumption we are exploring is for agents instead to forecast future real interest rates directly using AL. More general state-contingent AL forecast rules could also be investigated.

11 Summary and Conclusions

Using a nonlinear long-horizon stochastic New Keynesian model, which can have multiple steady states, we study the model dynamics generated by the interaction of output and inflation expectations and the realized values of these variables. Expectations play a key role in macroeconomic models, and we argue that the combination of AL (adaptive learning) and CD (cognitive discounting of more distant forecasts) provides a natural bounded-rationality framework for expectations in long-horizon models. Under AL, household and firm expectations of future inflation and output are formed and updated using statistical forecasting equations. In our long-horizon model, CD is motivated by the relative unreliability of econometric forecasts at long horizons due, e.g., to unanticipatable structural change. In our baseline calibration, which we examine in detail, there are two locally stable steady states: the steady state targeted by policy and a Great Depression steady state.

In order to incorporate both AL and CD, as well as exogenous autocorrelated productivity and mark-up shocks, we use a multiplicative formulation, which makes possible fast computation of key features of the model: number and location of steady states, associated phase diagrams generated using E-stabilty tools, showing which steady states are locally stable and providing their corresponding domains of attraction, and fast computations of time paths and phase diagrams. For initial expectations (π^e, y^e) the mean time paths of expected and actual (π, y) can be obtained. Net impulse response functions for exogenous random shocks are straightforward to compute. We find that stochastic shocks can occasionally lead to escapes from the target steady state to the Great Depression steady state unless they are offset by policy.

An important question then is whether, if the economy is in a Depression steady state (or in the domain of attraction of the depression steady state), there are policies that will push the economy to the steady state targeted by policy. Perhaps surprisingly, we find that a large anticipated temporary increase in government spending is not sufficient to push the economy into the domain of attraction of the target steady state. In contrast, policies that generate a large anticipated temporary increase in the price mark-up *can* be effective in returning the economy to the targeted steady state. We also find that a large anticipated temporary increase in government spending *can* be effective in pushing the economy out of Depression, and to the target steady state, *if* the composition of the additional government spending is on goods that have high substitutability with private consumption.

Appendix

A Mahler Series

The state dynamics of our model take the form the form

$$\begin{aligned} y_t &= g(1-\xi) + \Theta_{y1} \sum_{n\geq 1} \beta^n a_1^{m_1^n} a_2^{m_2^n} - \Theta_{y2} \sum_{n\geq 1} \beta^n a_1^{m_1^n} \\ Q(\pi_t) &= G(y_t) + \Theta_{\pi 1} \sum_{n\geq 1} \beta^n b^{l^n} + \Theta_{\pi 2} \sum_{n\geq 1} \beta^n \left(\frac{b^{l^n}}{b^{l^n} - (1-\xi) g/y^*} \right) \end{aligned}$$

The naive computational approach of truncating sums introduces numerical instability. Using log expansions, we transform the above dynamics into Mahler-type series, whose partial sums converge much faster than those under the naive approach.

In this section of the Appendix, we derive two Mahler-type expansions, denoted *bivariate* and *modified* Mahler series, and we state their generalized versions, whose derivations are straightforward extensions of those given below. Following the derivations is a brief discussion of the computational efficiency afforded by the expansions.

A.1 Mahler series

Lemma 1 (Bivariate Mahler) *Let* $0 < a, b, c < 1$, $x, y > 0$, *and*

$$F_{a,b,c}(x,y) = \sum_{n=1}^{\infty} a^n x^{b^n} y^{c^n} \tag{5}$$

Then the exact bivariate expansion of F *about* $(x,y) = (1,1)$ *is*

$$F_{a,b,c}(x,y) = \sum_{k,l\geq 0}^{\infty} \left(\frac{a b^k c^l}{1 - a b^k c^l} \right) \left(\frac{(\log x)^k (\log y)^l}{k!\, l!} \right) \tag{6}$$

Letting $\rho = \max\{|\log x| b, |\log y| c\}$, *and* R_M *the error associated with* $k + l \leq M$ *gives*

$$\left| R_M(x,y) \right| \leq \frac{a}{1-a} \frac{\rho^{M+1}}{(M+1)!} \exp(\rho). \tag{7}$$

Proof. The functional form of (6) suggests the proof strategy: change variables and exploit the global expansion of the exponential function. To this end, let $\hat{\star} = \log \star$. Then (1)

becomes

$$\begin{aligned}
F_{a,b,c}(x,y) &= \sum_{n=1}^{\infty} a^n \exp\left(b^n \cdot \hat{x}\right) \exp\left(c^n \cdot \hat{y}\right) \\
&= \sum_{n=1}^{\infty} a^n \left(\sum_{l \geq 0} \frac{\left(b^n \cdot \hat{x}\right)^l}{l!} \right) \left(\sum_{k \geq 0} \frac{\left(c^n \cdot \hat{y}\right)^k}{k!} \right) \\
&= \sum_{n=1}^{\infty} a^n \left(\sum_{l \geq 0, k \geq 0} \frac{(b^l)^n \cdot \hat{x}^l \cdot (c^k)^n \cdot \hat{y}^k}{l!k!} \right) \\
&= \sum_{l \geq 0, k \geq 0} \left(\sum_{n \geq 1} a^n \left(\frac{(b^l)^n \cdot \hat{x}^l \cdot (c^k)^n \cdot \hat{y}^k}{l!k!} \right) \right) \\
&= \sum_{l \geq 0, k \geq 0} \left(\frac{\hat{x}^l \cdot \hat{y}^k}{l!k!} \sum_{n \geq 1} a^n (b^l)^n (c^k)^n \right) \\
&= \sum_{l \geq 0, k \geq 0} \left(\frac{\hat{x}^l \cdot \hat{y}^k}{l!k!} \right) \left(\frac{a b^l \cdot c^k}{1 - a b^l \cdot c^k} \right) = \text{ equation (6)}
\end{aligned}$$

The error term follows from Taylor's theorem. ■

Lemma 2 (Modified Mahler) *Let* $0 < a, b < 1$, $x > 0$, *and* $0 < g < x^{b^n}$ *for* $n \geq 1$. *Set*

$$H_{a,b}(x) \;=\; \sum_{n \geq 1} a^n \left(\frac{x^{b^n}}{x^{b^n} - g} \right).$$

Then the exact expansion of H *about* $x = 1$ *is*

$$H_{a,b}(x) \;=\; \frac{a}{(1-a)(1-g)} + \sum_{k=1}^{\infty} \frac{(\log x)^k}{k!} \frac{a b^k}{1 - a b^k} \left[(-1)^k \frac{g}{(1-g)^{k+1}} \sum_{j=0}^{k-1} \left\langle {k \atop j} \right\rangle g^j \right].$$

The error term satisfies

$$\left| R_M(x) \right| \;\leq\; \frac{a \cdot g}{(1-a)(1-g)} \frac{\left(|\log x| b / (1-g) \right)^{M+1}}{(M+1)!} \exp\!\left(\tfrac{|\log x|}{1-b} \right).$$

Proof. Our assumptions imply $0 < g \cdot x^{-b^n} < 1$, so that

$$\frac{x^{b^n}}{x^{b^n} - g} = \frac{1}{1 - g \cdot x^{-b^n}} = \sum_{m \geq 0} \left(g \cdot x^{-b^n} \right)^m = \sum_{m \geq 0} g^m x^{-m b^n}.$$

It follows that

$$H_{a,b}(x) = \sum_{n\geq 1} a^n \left(\sum_{m\geq 0} g^m x^{-mb^n} \right)$$
$$= \sum_{m\geq 0} g^m \left(\sum_{n\geq 1} a^n (x^{-m})^{b^n} \right).$$

By the standard Mahler expansion,

$$\sum_{n\geq 1} a^n (x^{-m})^{b^n} = \sum_{k\geq 0} \frac{(-m\log x)^k}{k!} \cdot \frac{ab^k}{1-ab^k}.$$

Thus

$$H_{a,b}(x) = \sum_{m\geq 0} g^m \sum_{k\geq 0} \frac{(-m\log x)^k}{k!} \cdot \frac{ab^k}{1-ab^k}$$
$$= \frac{a}{(1-a)(1-g)} + \sum_{k\geq 1} \frac{(\log x)^k}{k!} \frac{ab^k}{1-ab^k} \left(\sum_{m\geq 0} g^m (-m)^k \right).$$

Now recall

$$\sum_{m\geq 0} g^m (-m)^k = (-1)^k \frac{g}{(1-g)^{k+1}} \sum_{j=0}^{k-1} \left\langle {k \atop j} \right\rangle g^j.$$

Substituting gives the claimed expansion. The error bound follows from the standard remainder estimate for the Mahler expansion, together with the uniform bound on the geometric prefactor. ■

Below we state generalized versions of the above lemmas. The proofs are routine generalizations and thus omitted.

Lemma 3 (Generalized Mahler Expansion) *Let $0 < \beta, m_1, \ldots, m_N < 1$, and let $x_1, \ldots, x_N > 0$. Define*

$$F_{\beta,\mathbf{m}}(\mathbf{x}) := \sum_{n=1}^{\infty} \beta^n \prod_{s=1}^{N} x_s^{m_s^n}.$$

Then the exact expansion of F about $\mathbf{x} = (1, \ldots, 1)$ is

$$F_{\beta,\mathbf{m}}(\mathbf{x}) = \sum_{k_1,\ldots,k_N \geq 0} \left(\frac{\beta \cdot m_1^{k_1} \cdots m_N^{k_N}}{1 - \beta \cdot m_1^{k_1} \cdots m_N^{k_N}} \right) \prod_{s=1}^{N} \frac{(\log x_s)^{k_s}}{k_s!}.$$

Letting $\rho = \max_s\{|\log x_s| m_s\}$, the error from truncating to total degree $\sum k_s \leq M$ satisfies

$$|R_M(\mathbf{x})| \leq \frac{\beta}{1-\beta} \cdot \frac{\rho^{M+1}}{(M+1)!} \cdot \exp(\rho).$$

Lemma 4 (Generalized Modified Mahler Expansion) *Let* $0 < \beta, l_1, \ldots, l_N < 1$, $x_1, \ldots, x_N > 0$, *and* $\delta > 0$ *such that for all* $n \geq 1$,

$$\prod_{s=1}^{N-1} x_s^{l_s^n} > \delta \, .$$

Define

$$H_{\beta,\mathbf{l}}(\mathbf{x}) := \sum_{n=1}^{\infty} \beta^n \cdot \frac{\prod_{s=1}^{N} x_s^{l_s^n}}{\prod_{s=1}^{N-1} x_s^{l_s^n} - \delta} \, .$$

Then the exact expansion about $\mathbf{x} = (1, \ldots, 1)$ *is*

$$H_{\beta,\mathbf{l}}(\mathbf{x}) = \frac{\beta}{(1-\beta)(1-\delta)} + \sum_{m=1}^{\infty} \sum_{\substack{k_1+\cdots+k_N=m \\ k_s \geq 0}} \left(\prod_{s=1}^{N} \frac{(\log x_s)^{k_s}}{k_s!} \right) \cdot \frac{\beta \cdot \prod_{s=1}^{N} l_s^{k_s}}{1 - \beta \cdot \prod_{s=1}^{N} l_s^{k_s}} \cdot \frac{\delta}{(1-\delta)^{m+1}} \sum_{r=0}^{m} \left\langle {m \atop r} \right\rangle \delta^r \, .$$

Letting $\rho = \max_s \{ |\log x_s| l_s \}$, *the remainder from truncating total degree* $m > M$ *satisfies*

$$|R_M(\mathbf{x})| \leq \frac{\beta\delta}{(1-\beta)(1-\delta)} \cdot \frac{\rho^{M+1}}{(M+1)!} \cdot \exp\left(\sum_{s=1}^{N} \frac{|\log x_s|}{1 - l_s} \right) \, .$$

A.2 Computational efficiency of Mahler series

The expansions above are generalizations of the hypergeometric series

$$H_{a,b}(x) = \sum_{n=1}^{\infty} a^n x^{b^n}, \text{ with } 0 < a, b < 1 \text{ and } x \in (1-\delta, 1+\delta), \tag{8}$$

where, for our purposes, $a = 0.99$ (and, so, near unity), $b \in (0.75, 0.95)$ and $0 < \delta < 1/4$, though none of the results here turn qualitatively on these ranges.

There is no closed form representation of this series, so numerical work requires approximation methods. Unfortunately, the naive method of truncation at M is not very efficient, and can introduce numerical instabilities for M larger than ≈ 500, depending on values of a, b and δ.

To make improvements, first note that H is real-analytic on $(0, \infty)$ – in fact, it can be analytically continued to $\mathbb{C} \setminus \{0\}$ and is characterized by an associated functional equation, though we will not need these features. Because it is real analytic, we can expand H around the natural base-point $x = 1$; however, as noted in the previous section, the multiplicative structure suggests a "log-expansion" would be more natural. It can be shown, using the techniques provided in the proofs above, that

$$H_{a,b}(x) = \sum_{k \geq 0}^{\infty} \left(\frac{ab^k}{1 - ab^k} \right) \left(\frac{(\log x)^k}{k!} \right). \tag{9}$$

This expansion is global in the sense that it holds for all $x > 0$, and truncation of this series provides a very efficient approximation of H.

So what gives? Here are some reasons for the improved efficiency:

- The exponents in (9) are integral whereas those in (8) are floating point and thus more expensive to compute.
- The RHS of (9) is a Taylor expansion, which are generally very efficient, particularly for real-analytic functions.
- Expanding in $\log x$ rather that x provides coefficients with more rapid decay.

There is an explicit metric that can be used to show that (9) is far superior to (8). We can compute the amount of new information, in the Kolmogorov/Tikhomirov sense, that each additional term provides. More specifically, we measure the *information content*, $\mathscr{I}$, as the minimum number of binary digits of accuracy obtained by a given truncation level M:

$$\mathscr{I}(M) \;\approx\; -\log_2\big(\text{uniform error bound of } M \text{ truncation}\big).$$

Note that larger $\mathscr{I}$ corresponds to a truncation that provides more information about the function values, i.e. higher accuracy.

This notion can be linked to more intuitive notions of information: if a real number x is known to be between zero and one then boolean logic allows for the determination if it is above or below $1/2$ – one question, one bit of information. To determine if it is in some quarter interval, we first determine if it is above or below $1/2$, and then if it is above or below the midpoint of the interval just determined – two questions, two bits of information. Thus to know if x belongs to an interval of length ε requires $-\log_2 \varepsilon$ assessments, i.e. bits of information.

Suppressing the subscripts a, b, write the truncations as

$$\begin{aligned} H_{\text{geom}}(x,M) &= \sum_{n=1}^{M} a^n x^{b^n} \\ H_{\log}(x,M) &= \sum_{k=0}^{M-1} \left(\frac{ab^k}{1-ab^k}\right)\left(\frac{(\log x)^k}{k!}\right), \end{aligned}$$

where the index ranges are chosen to ensure the same number of summands. Then, on $(1-\delta, 1+\delta)$,

$$\mathscr{I}_{\text{geom}}(M) = \;-\log_2 \|H - H_{\text{geom}}(\,\cdot\,,M)\|_\infty \approx (M+1)\,(-\log_2 a) \;-\log_2(1-a) \;-\tfrac{\delta}{\log 2}\, b^{M+1}$$

$$\mathscr{I}_{\log}(M) = \;-\log_2 \|H - H_{\log}(\,\cdot\,,M)\|_\infty \approx \sum_{m=0}^{M-1} \log_2\!\Big(\tfrac{m+1}{\delta b}\Big) \;-\log_2\!\Big(\tfrac{\delta b}{1-a}\Big) \;-\tfrac{\delta b}{\log 2}$$

Using the values $a = 0.99$, $b = 0.9$, and $\delta = 0.25$, which are consistent with the applications of these results to our modeling framework, with $M = 10$,

$$\mathscr{I}_{\text{geom}}(229 \cdot M) < \mathscr{I}_{\log}(M) < \mathscr{I}_{\text{geom}}(230 \cdot M).$$

Thus over 2000 terms of the original series are needed to provide the same information content as 10 terms of the transformed series. This explains the two orders-of-magnitude decrease in computational time needed to approximate the our model.

B Fiscal policy

In this section we extend the modified model to allow for anticipated changes to fiscal policy. The particular experiment we have in mind is as follows: in period t government spending is changed to $g + \Delta g$, and it is announced that this change will last exactly T periods, so that in period $t + T$ government spending reverts, permanently, to g. It is assumed that agents understand this policy, and incorporate it into their decision making.

Under steady state learning, and with general fiscal policy $\{g_t\}$, the IS and AS relations are given by

$$y_t = (1-\xi)g_t + \frac{1-\beta}{\beta} \sum_{n \geq 1} \beta^n D^e_{t+n} \left(y^e_{t+n} - (1-\xi) g^e_{t+n} \right) \tag{10}$$

$$\phi'(\pi_t)\pi_t = G(y_t, g_t) + \sum_{n \geq 1} \beta^n G\left(y^e_{t+n}, g^e_{t+n} \right), \tag{11}$$

where G and D^e_{t+n} are given below.

$$G(y) = \frac{\nu}{\alpha} \left(\frac{y}{A} \right)^{\frac{1+\varepsilon}{\alpha}} - \frac{y(\nu-1)}{y-(1-\xi)g} \equiv \gamma \cdot y^{\frac{1+\varepsilon}{\alpha}} - \frac{y(\nu-1)}{y-(1-\xi)}$$

$$\hat{D}_n(\pi^e, y^e) = \beta^n \left(\left(\frac{r^* \pi^e}{R(\pi^e, y^e)} \right)^{\frac{m_r}{1-m_r}} \right)^{1-m_r^n}.$$

Now assume that in period t, government spending is given by $\hat{g} = g + \Delta g$, and that in period $t + T$, government spending returns to g; and further assume that agents know this.

Letting $\Delta G(y_t, g, \hat{g}) = G(y, \hat{g}) - G(y, g)$, we have

$$\begin{aligned}
y_t &= (1-\xi)\hat{g} + \frac{1-\beta}{\beta}\sum_{n=1}^{T-1}\beta^n D^e_{t+n}\left(y^e_{t+n} - (1-\xi)\hat{g}\right) + \frac{1-\beta}{\beta}\sum_{n\geq T}\beta^n D^e_{t+n}\left(y^e_{t+n} - (1-\xi)g\right) \\
&= (1-\xi)g + (1-\xi)\Delta g + \frac{1-\beta}{\beta}\sum_{n\geq 1}\beta^n D^e_{t+n}\left(y^e_{t+n} - (1-\xi)g\right) + \frac{1-\beta}{\beta}\sum_{n=1}^{T-1} D^e_{t+n}(1-\xi)(g-\hat{g}) \\
&= T_y\left(\pi^e_{t-1}, y^e_{t-1}, g\right) + (1-\xi)\Delta g\left(1 - \frac{1-\beta}{\beta}\sum_{n=1}^{T-1} D^e_{t+n}\right) \\
&\equiv T_y\left(\pi^e_{t-1}, y^e_{t-1}\right) + \Omega_y\left(\pi^e_{t-1}, y^e_{t-1}, \hat{g}, g, T\right) \\
\phi'(\pi_t)\pi_t &= G(y_t, \hat{g}) + \sum_{n=1}^{T-1}\beta^n G(y_{t+n}, \hat{g}) + \sum_{n\geq T}\beta^n G\left(y^e_{t+n}, g\right) \\
&= S\left(\pi^e_{t-1}, y^e_{t-1}, g\right) + \Delta G(y_t, g, \hat{g}) + \sum_{n=1}^{T-1}\beta^n \Delta G(y_{t+n}, \hat{g}, g) \\
&\equiv S\left(\pi^e_{t-1}, y^e_{t-1}, g\right) + \Delta G(y_t, g, \hat{g}) + \Omega_\pi\left(\pi^e_{t-1}, y^e_{t-1}, \hat{g}, g, T\right)
\end{aligned}$$

It follows that dynamics under anticipated policy change can be determined using the usual, computationally efficient maps T_y and S together with computationally inexpensive terms $\Omega_\star$.

C Anticipated productivity shocks

Here we consider anticipated supply-chain bottlenecks, proxied by known changes to the time-path of productivity, A. Noting that A impinges on dynamics only through G, see equation (11), we follow the reasoning and notation of the policy shocks section by writing

$$\Delta G^A(y) = G\left(y, \hat{A}\right) - G(y, A),$$

where A is the calibrated value of productivity and $\hat{A}$ is the bottleneck proxy; also, here and below, the superscript A indicates that these particular notations refer to anticipated changes in productivity.

Now assume that in period t_0 it is "announced" that in periods $t_0 + n$ for $n = \underline{T}, \ldots, \bar{T}$. As a benchmark, take $\underline{T} = 2$ and $\bar{T} = 8$, where units are quarters. Define

$$\Omega^A_t\left(\pi^e_{t-1}, y^e_{t-1}\right) = \chi\left(t_0 + \underline{T} \leq t \leq t_0 + \bar{T}\right) \cdot \Delta G^A\left(T_y\left(\pi^e_{t-1}, y^e_{t-1}\right)\right) + \sum_{n=1}^{\bar{T}+t_0-t}\beta^n \Delta G^A\left(y^e_{t+n}\right),$$

where χ is the characteristic function. Letting

$$S\left(\pi^e_{t-1}, y^e_{t-1}\right) = G\left(T_y\left(\pi^e_{t-1}, y^e_{t-1}\right)\right) + \sum_{n\geq 1}\beta^n G\left(y^e_{t+n},\right)$$

we get the AS relation

$$\phi'(\pi_t)\pi_t = S\left(\pi^e_{t-1}, y^e_{t-1}\right) + \Omega^A_t\left(\pi^e_{t-1}, y^e_{t-1}\right), \tag{12}$$

which can be joined to the IS relation (10) to determine the attendant dynamics.

D NK with AL forecasts of real rates

Consider the canonical NK model with white noise demand shock and forward PR:

$$y_t = \hat{E}_t y_{t+1} - \sigma \hat{E}_t r_{t+1} + g_t \tag{13}$$
$$\pi_t = \beta E_t \pi_{t+1} + \lambda y_t \tag{14}$$
$$r_t = i_{t-1} - \pi_t \tag{15}$$
$$i_t = \phi_y \hat{E}_t y_{t+1} + \phi_\pi \hat{E}_t \pi_{t+1} \tag{16}$$

D.1 The standard approach

The standard approach steps equation (15) forward one period and substitutes it and the policy rule (16) into the IS relation (13), thus eliminating the need to forecast real rates. The attendant model, which is purely forward-looking, is given as

$$y_t = (1-\sigma\phi_y)\hat{E}_t y_{t+1} + \sigma(1-\phi_\pi)\hat{E}_t \pi_{t+1} + g_t$$
$$\pi_t = \lambda(1-\sigma\phi_y)\hat{E}_t y_{t+1} + (\beta + \lambda\sigma(1-\phi_\pi))\hat{E}_t \pi_{t+1} + \lambda g_t$$

Letting $x_t = (y_t, \pi_t)'$ and providing PLM $x_t = x^e$ we get the ALM

$$x_t = \underbrace{\begin{pmatrix} 1-\sigma\phi_y & \sigma(1-\phi_\pi) \\ \lambda(1-\sigma\phi_y) & \beta + \lambda\sigma(1-\phi_\pi) \end{pmatrix} x^e}_{T(x^e)} + \begin{pmatrix} 1 \\ \lambda \end{pmatrix} g_t.$$

Writing $T(x^e) = \mathscr{T} \cdot x^e$, we have a classic result:

Claim: If $\phi_y = 0$ then $\phi_\pi > 1$ is necessary and sufficient for E-stability.[5]

Proof. Setting $\phi_y = 0$ we find that

$$\mathrm{Det}(\mathscr{T} - I_2) = -\sigma\lambda(1-\phi_\pi) \text{ and } \mathrm{Tr}(\mathscr{T} - I_2) = \beta - 1 - \mathrm{Det}(\mathscr{T} - I_2).$$

Thus $\mathrm{Det}(\mathscr{T} - I_2) > 0 \iff \phi_\pi > 1$; the trace formula then shows that $\mathrm{Det}(\mathscr{T} - I_2) > 0 \implies \mathrm{Tr}(\mathscr{T} - I_2) < 0$. ■

[5] Assumes, naturally, that $\sigma, \lambda > 0$ and $0 < \beta < 1$.

E Bifurcation analysis

We have the following computational observation: for some calibrations there are at least two steady states (often at least three), including the target steady state, and all observed steady states are on the same component of the $\dot{\pi} = 0$ nullcline; for other calibrations there are at least two steady states (often at least three), however the target steady state and the lower steady state(s) are on distinct components of the $\dot{\pi} = 0$ nullcline. Additionally, in the former, "single component" case the $\dot{\pi} = 0$ nullcline crosses the $\dot{y} = 0$ nullcline, at the target steady state, from below, whereas in the latter case the crossing is from above.

How to make sense of these observations? Imagine starting in the single-component case and then changing the calibration smoothly and observing the resulting deformation of the $\dot{\pi} = 0$ nullcline. It would seem that, as a particular calibration (or collection of calibrations) is approached, the $\dot{\pi} = 0$ nullcline becomes vertical at the target steady state. Crossing this singularity somehow results the connected component splitting in two.

To examine this further, adopt the following notation: $\hat{\phi}(\pi) = \phi'(\pi)\pi$, so that the $\dot{\pi} = 0$ nullcline becomes

$$\hat{\phi}(\pi) = G(T_y(\pi, y)) + \sum_{n \geq 1} \beta^n G\left(y^* \left(\frac{y}{y^*}\right)^{m_y^n}\right). \tag{17}$$

Differentiating, we get

$$\left.\frac{dy}{d\pi}\right|_{y=y^*,\pi=\pi^*} = \frac{\hat{\phi}(\pi^*) - G'(y^*)\left.\frac{\partial}{\partial \pi} T_y(\pi, y)\right|_{y=y^*,\pi=\pi^*}}{G'(y^*)\left(\left.\frac{\partial}{\partial y} T_y(\pi, y)\right|_{y=y^*,\pi=\pi^*} + \frac{\beta m_y}{1-\beta m_y}\right)} \tag{18}$$

The numerator and $G'(y^*)$ are not problematic, so the singularity occurs when

$$\left.\frac{\partial}{\partial y} T_y(\pi, y)\right|_{y=y^*,\pi=\pi^*} = -\frac{\beta m_y}{1-\beta m_y}. \tag{19}$$

Intuitively, when y^e rises there are competing pressures on the Phillips curve. The rise in y^e raises expected future marginal costs, putting upward pressure on inflation; however the rise in y^e also raises real rates via the Taylor rule, thus depressing inflation. When these pressures balance, a rise in y^e has no impact on inflation, leading to a $\dot{\pi} = 0$ nullcline with vertical slope.

Computation of the LHS of (19) is straightforward, yielding

$$\left.\frac{\partial}{\partial y} T_y(\pi, y)\right|_{y=y^*,\pi=\pi^*} = (y^* - (1-\xi)g)\phi_y \left(\frac{m_r}{m_r - 1}\right)\left(\frac{R^* - 1}{\pi^* r^* y^*}\right)\left[1 + \left(\frac{1-\beta}{\beta}\right)\left(\frac{\beta m_r}{1-\beta m_r}\right)\right] + \left(\frac{1-\beta}{\beta}\right)\left(\frac{\beta m_y}{1-\beta m_y}\right). \tag{20}$$

Using (19) and (20), we can determine, for example, pairs (m_r, m_y) that will induce a two-component nullcline: see Figure below.

F Linearized model and fiscal policy

Recall that, in our model, simulations of announced, temporary fiscal expansion resulted in a negative impact effect on inflation. This seems, ex ante, incongruous with the standard Keynesian narrative that deficits induce inflation. Of course, there is a lot to unpack here: we don't have deficits – nor does the standard NK model – and most analysis of fiscal multipliers have focused on output and unemployment.

Below we examine the linearized version of our model. With find that, in our model, increased spending is *always* deflationary; however, if the reduced form parameters are allowed to depart from the values associated with our deep parameters there are inflationary calibrations.

I think the "incongruity" mentioned above arises from the definition of "output gap" being employed. Here I have log-linearized around the non-stochastic steady state; however, Woodford often presents the gap as the discrepancy between output and its flexible price analog, which, itself, would obviously depend on government spending. In his models, I think this dependence subsumes the direct effect of spending on inflation in the Phillips curve.[6]

The linearized model

Make the simplifying assumptions that $\xi, \zeta = 0$ and $A, \alpha = 1$. Letting $\theta = \bar{g}/y^*$, I get (all variables are in proportional deviation from steady state)

$$\begin{aligned} y_t &= E_t y_{t+1} - (1-\theta)(R_t - E_t\pi_{t+1}) + \theta(g_t - E_t g_{t+1}) \\ \pi_t &= \beta E_t \pi_{t+1} + \kappa y_t - \gamma g_t \\ R_t &= \phi_\pi E_t \pi_{t+1} + \phi_y E_t y_{t+1}, \text{ where} \end{aligned} \tag{21}$$

$$\gamma = \left(\frac{\nu - 1}{(\pi^*)^2 \varphi}\right)\left(\frac{\theta}{(1-\theta)^2}\right), \quad \kappa = \left(\frac{2-\theta}{\theta}\right)\gamma, \quad \phi_\pi = B, \text{ and } \phi_y = \varphi_y\left(\frac{R^* - 1}{R^*}\right).$$

For the moment allowing κ and γ to be arbitrary up to satisfying $0 < \gamma < \kappa < 1$ (i.e. abandoning the attachment to the deep parameters, but retaining relative sizes and reasonable values), with the help of Mathematica, I find that the REE coefficient of inflation on gov-

[6] On page 461 of his tome he refers to the potential for spending to act as a negative cost-push shock, and this occurs when the optimal output level does not coincide with the flexible price level. I think this feature requires monetary policy working to enforce the optimal output level.

ernment spending is positive when

$$0 < \gamma < \left(\frac{\theta}{\phi_y (1-\theta) \left(\frac{\rho}{1-\rho} \right) + 1} \right) \kappa \equiv \bar{\gamma} \kappa, \tag{22}$$

where the last equality defines $\bar{\gamma} < \theta$. Using the relation from our model that $\kappa = \left(\frac{2-\theta}{\theta} \right) \gamma$, we conclude fiscal policy is expansionary if and only if

$$\phi_y \left(\frac{\rho}{1-\rho} \right) < 1. \tag{23}$$

Thus the interest-rate response to output must not be too large relative to the persistence of the shock.

Summary:

- As seen in the Phillips curve (21), the direct impact of government spending on inflation is negative (i.e. $\gamma > 0$). This reflects the marginal cost channel analogous to the RBC model.
- For the equilibrium impact of spending to be inflationary, this negative impact must be small relative to the aggregate demand effect, captured by part of κ:[7] according to (22), $\frac{\gamma}{\kappa} \leq \bar{\gamma} < \theta$.
- In our model, under the baseline calibration, except with $\xi, \zeta = 0$ and $A, \alpha = 1$, we have that $\phi_y = \varphi_y \left(\frac{R^* - 1}{R^*} \right) = 0.122$, whereas $\frac{1-\rho}{\rho} = 0.111$: thus (23) is violated and fiscal policy is not inflationary. However, a small reduction in either ϕ_y or ρ will flip this result.

References

BENHABIB, J., G. W. EVANS, AND S. HONKAPOHJA (2014): "Liquidity Traps and Expectation Dynamics: Fiscal Stimulus or Fiscal Austerity?," *Journal of Economic Dynamics and Control*, 45, 220–238.

BRANCH, W. A., AND G. W. EVANS (2006): "A Simple Recursive Forecasting Model," *Economic Letters*, 91, 158–166.

CASTLE, J. L., M. P. CLEMENTS, AND D. F. HENDRY (2019): *Forecasting: An Essential Introduction*. Yale University Press, New Haven.

[7] The parameter κ also captures some of the marginal cost.

CHRISTIANO, L., AND M. EICHENBAUM (1992): "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review*, 82, 430–450.

EGGERTSSON, G. B. (2008): "Great Expectations and the End of the Depression," *American Economic Review*, 102, 476–1516.

——— (2012): "Was the New Deal Contractionary?," *American Economic Review*, 98, 524–555.

EUSEPI, S. (2010): "Central Bank Communication and the Liquidity Trap," *Journal of Money, Credit and Banking*, 42, 373–398.

EUSEPI, S., AND B. PRESTON (2010): "Central Bank Communication and Expectations Stabilization," *American Economic Journal: Macroeconomics*, 2, 235–271.

——— (2011): "Expectations, Learning and Business Cycle Fluctuations," *American Economic Review*, 101, 2844–2872.

——— (2018): "The Science of Monetary Policy: An Imperfect Knowledge Perspective," *Journal of Economic Literature*, 56, 3–59.

EVANS, G. W., AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, New Jersey.

EVANS, G. W., S. HONKAPOHJA, AND K. MITRA (2009): "Anticipated Fiscal Policy and Learning," *Journal of Monetary Economics*, 56, 930–953.

——— (2022): "Expectations, Stagnation and Fiscal Policy: a Nonlinear Analysis," *International Economic Review*, 63, 1397–1425.

EVANS, G. W., AND B. MCGOUGH (2020): "Adaptive Learning in Macroeconomics," *Oxford Research Encyclopedia of Economics and Finance*, https://doi.org/10.1093/acrefore/9780190625979.013.508.

GABAIX, X. (2020): "A Behavioral New Keynesian Model," *American Economic Review*, 110, 2271–2327.

HENDRY, D. F., AND B. NIELSEN (2007): *Econometric Modeling*. Princeton University Press, Princeton, New Jersey.

HENDRY, D. F., AND F. PRETIS (2016): "All Change! : The Implications of Non-stationarity for Economic Modelling, Forecasting and Policy," Occasional paper, Oxford Martin School Policy Paper.

MILANI, F. (2007): "Expectations, Learning and Macroeconomic Persistence," *Journal of Monetary Economics*, 54, 2065–2082.

ORPHANIDES, A., AND J. C. WILLIAMS (2002): "Robust Monetary Policy with Unknown Natural Rates," *Brookings papers on Economic Activity*, 2, 63–118.

——— (2005): "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning and Expectations," *Journal of Economic Dynamics and Control*, 29, 1927–1950.

PRESTON, B. (2005): "Learning about Monetary Policy Rules when Long-Horizon Expectations Matter," *International Journal of Central Banking*, 1, 81–126.