The New Keynesian Price Puzzle: Reinterpreting Inflation Dynamics

Neil White*
Amherst College

January 7, 2025

For the most recent version of this paper, visit
`https://sites.google.com/view/neilwhite/research`

Abstract

The "price puzzle"—a short-run increase in prices in response to higher interest rates—is a counterintuitive but common feature of estimated monetary policy shocks. I show that it is also a clear implication of rational expectations in the textbook New Keynesian model: equilibria consistent with common identifying assumptions and estimates of interest rate responses almost always result in a price puzzle. Nevertheless, offsetting the short-run inflationary effects of expansionary fiscal shocks in the model often calls for higher interest rates. These results illustrate the pitfalls of inferring causal effects of *endogenous* policy changes from responses to *exogenous* policy shocks.

Keywords: Monetary policy, fiscal policy.

JEL classification:

*Amherst College, Department of Economics. 100 Boltwood Ave., Converse Hall, Amherst, MA 01002 U.S.A. Email: `nwhite@amherst.edu`. For helpful comments and suggestions, I thank Jake Blackwood, Gabe Chodorow-Reich, Junkyoo Choi, John Cochrane, Niklas Kroner, and seminar participants at the Macro Dynamics workshop and SEA. I am grateful for funding from the Provost's Research Fellowship from Amherst College that supported this research.

1 Introduction

Figure 1 displays prominent examples of the "price puzzle": estimated monetary policy shocks that increase short-term nominal interest rates (second column) lead to a temporarily *rising* price level (first column) that does not begin to fall until months or even years after the shock. It is a puzzle because it goes against the almost universally held intuition that higher interest rates should lower inflation in the short run, not raise it. Reflecting this strong prior belief, macroeconometricians have proposed many fixes to explain away the price puzzle, with most implying that its presence is a sign of poor identification.

But given its pervasiveness in empirical work, maybe the price puzzle reflects something true about the macroeconomy—something important to understand, rather than to be explained away. So, I take a different approach; instead of looking for a new empirical patch, I explore theoretical explanations of the price puzzle. More precisely, I ask if there is a simple model in which the paths of nominal interest rates from the second column of Figure 1 result in price-level responses similar to those in the first column.

The answer is yes, and it is the simplest version of the textbook[1] sticky-price New Keynesian model. The model consists of a linearized consumption Euler equation expressed in terms of the output gap and a forward-looking Phillips curve that characterizes the linearized solution to monopolistically competitive firms' profit maximization problems when prices are sticky:

$$x_t = \mathbb{E}_t x_{t+1} - \sigma(i_t - \mathbb{E}_t \pi_{t+1}) \tag{1}$$

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t, \tag{2}$$

where i_t is the nominal interest rate, x_t is the output gap, π_t is inflation, and $\mathbb{E}_t$ is the rational-expectations operator conditional on time-t information. All variables are expressed in deviations from steady state.[2]

To see why this model produces price-level responses to those in the first column of Figure 1, let's ask what happens to the price level when policymakers in this model set the same interest rates as in the second column of Figure 1. Let's also make the additional assumption that the impact effect on inflation in the model is exactly the same as the effect estimated in the data.[3] This assumption is sufficient to pin down a unique equilibrium in the model.

[1] Prominent examples include Woodford (2003) and Galí (2015).

[2] The parameters σ and β reflect household preferences, and κ primarily captures the degree of price stickiness. I describe the ingredients in more detail in Section 3.

[3] This value is zero by the identifying assumptions in all except Gertler and Karadi (2015).

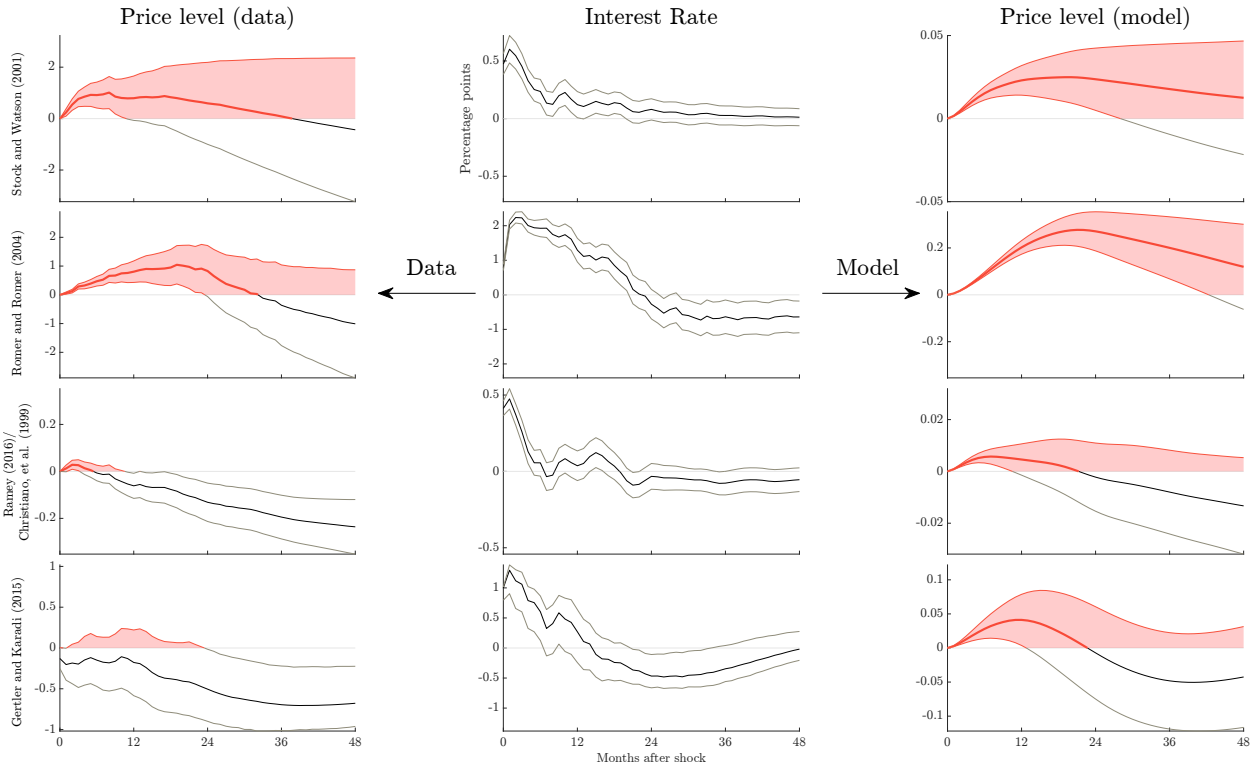


FIGURE 1: EXAMPLES OF THE PRICE PUZZLE

Notes: Empirical response of the price level (left column) and the short-term interest rate (middle column) to contractionary monetary policy shocks. Response of the price level (right column) in the New Keyensian model to the interest rate path in the middle column, assuming no contemporaneous response of inflation with $\sigma = 1$, $\kappa = 0.1717$, and $\beta = 0.99$ from Galí (2015). First row: Stock and Watson (2001); second row: Romer and Romer (2004); third row: Ramey's (2016) monthly version of Christiano et al. (1999); fourth row: Gertler and Karadi (2015). Red areas in the outer columns indicate *increases* in the price level in response to contractionary monetary policy shocks. The empirical price level is the log of the consumer price index, and the short-term interest rate is the one-year Treasury rate for Gertler and Karadi (2015) and the federal funds rate for all others.

The results are displayed in the last column of Figure 1.

An off-the-shelf calibration[4] of a simple *two-equation* model produces impulse responses that exhibit a price puzzle qualitatively similar to their empirical counterparts.[5] To be clear, I emphasize the similarity is *qualitative* and its degree differs across estimates; although its small scale is appealing, the two-equation model is probably not rich enough to give a

[4]From chapter 2 of Galí's (2015) textbook. In this calibration, $\beta = 0.99$, $\sigma = 1$, and $\kappa = 0.1717$. I assume agents in the model have perfect knowledge of the sequence $\{i_t\}_{t \geq 0}$. The price level response is the cumulative response of period-to-period inflation responses. To produce the error bands consistent with the empirical ones, I feed in each interest rate response from bootstrap estimates separately.

[5]By contrast, the model in Smets and Wouters (2007) contains 37(!) equations and its price-level responses to a monetary policy shock are still substantially different from those above and never rise above zero.

quantitative match. Nor, however, have I attempted any estimation exercise to match the data, but rather I have simply taken a standard calibration.[6]

The right column of Figure 1 is the first main result of the paper: the textbook New Keynesian model predicts that the path of nominal interest rates from estimates of the effects of monetary policy shocks should lead to a price puzzle in equilibria consistent with assumptions made in the construction of those estimates. Mathematically, the price puzzle is no puzzle at all; it is exactly what textbook models predict will happen after monetary policy shocks that "look like" their empirical counterparts. It is perhaps a previously unrecognized strength of the New Keynesian model that it can reproduce such a stubbornly robust empirical result.

The remainder of the paper is devoted to exploring this result in greater detail: its robustness, the intuition behind it, and its implications for the conduct of monetary policy.

It is tempting to immediately jump in to discuss the specific behavior of economic agents within the model that produces the price puzzle—an important discussion, to be sure, but one of secondary importance. Because it is an *equilibrium* of the New Keynesian model, the price puzzle is not confusing at all to the people living in that world. The price-level responses in Figure 1 are only puzzling to economists living in the real world outside simple the New Keynesian model. Therefore, I first explore what leads the New Keynesian model's predictions to differ so much from the conventional intuition that higher interest rates should, in the short run, lower inflation not raise it.

To that end, I show that the New Keynesian model can be expressed as the sum (or average) of two simpler models[7]. Each of these models consist of: 1) a static IS curve that posits that the output gap is inversely proportional to the *ex ante* real interest rate:

$$x_t = -\sigma(i_t - \pi_t^e),$$

and 2) a dynamic Phillips curve in which current inflation depends on inflation expectations and the output gap:

$$\pi_t = \pi_t^e + \kappa x_t.$$

[6] I show below that the shape of the responses are largely insensitive to most model parameters.

[7] These are studied in slightly different context in Cochrane (2023, 2024)

The two models differ in how expectations are formed. In the simple "Adaptive Expectations" model, expected inflation (π_t^e) in both equations is simply inflation from the previous period (π_{t-1}). In the simple "Rational Expectations" model, expected inflation is (unsurprisingly) the rational expectations solution to variables ($\mathbb{E}_t\pi_{t+1}$).

The "Adaptive Expectations" (**AE**) model is identical to the many models from the earlier New Keynesian literature and currently popular intermediately macroecononomics textbooks[8]. Accept for a moment these models reflect the conventional intuition of how monetary policy affects the real economy. Then the difference between this model and the full New Keynesian (**NK**) model is *the same* structural model as the intermediate macro model but with *rational* expectations (**RE**). Put simply:

$$\textbf{NK} = \textbf{AE} + \textbf{RE}.$$

This result, demonstrated in more detail in section 2, motivates a straightforward decomposition of the responses of prices in Figure 1. Given a response of inflation or the price level to monetary policy shock in the New Keynesian model, one can reasonably decompose that response into effects from the conventional wisdom plus that same model under rational expectations. Using this method, I show that price puzzle's unintuitiveness can be explain almost entirely by rational expectations: when the **NK** model exhibits a price puzzle, it is almost entirely accounted for by the **RE** model.

2 A New Keynesian model-based decomposition

In this section I describe the solution to the New Keynesian model in both discrete and continuous time, describe in detail the construction of the model-based impulse responses in Figure 1, and the decomposition of these responses into adaptive- versus rational expectations model responses.

2.1 Solutions to the New Keynesian model

The non-policy block of the linearized New Keynesian model from the introduction, as derived from microfoundations in Woodford (2003) and Galí (2015), is repeated here for

[8]These include Hall and Mankiw (1994), Taylor (1994, 1999), and Hetzel (2000), as well as the textbook intermediate macroeconomic models described in, e.g., Jones (2021) and Mankiw (2022).

convenience:

$$x_t = \mathbb{E}_t x_{t+1} - \sigma(i_t - \mathbb{E}_t \pi_{t+1})$$
$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t,$$

where, again, π_t is inflation, x_t is the output gap, and i_t is the nominal interest rate, all expressed in deviations from steady state, and the positive parameters β, σ, and κ are standard and reflect the discount factor, households' intertemporal elasticity of substitution, and the degree of price stickiness, respectively. The equivalent continuous-time version of the model is

$$\frac{1}{dt}\mathbb{E}_t[dx(t)] = \sigma(i(t) - \pi(t)) \tag{3}$$
$$\frac{1}{dt}\mathbb{E}_t[d\pi(t)] = \rho\pi(t) - \kappa x(t), \tag{4}$$

where $\rho = \beta^{-1} - 1$ is the instantaneous discount rate. The continuous-time model is especially useful for deriving some of the results below. Typically, disturbance terms meant to capture aggregate demand and supply shocks are appended to the end of each expression. I simply omit them since the focus for now is on nominal interest rate shocks.

As shown in Cochrane (2023), the solutions to these models can be expressed as two-sided moving averages of past and future nominal interest rates plus an additional condition. Throughout the remainder of the paper, I assume that time begins at $t = 0$ in the steady state with the announcement of an interest rate path for $t \geq 0$ and, where applicable, agents in the model have perfect foresight thereafter. The discrete-time solution is given by

$$\pi_{t+1} = \frac{\kappa\sigma}{\mu_1 - \mu_2}\left\{C\mu_1^{-t} + \sum_{j=0}^{t}\mu_1^{-j} i_{t-j} + \sum_{j=1}^{\infty}\mu_2^{j}\mathbb{E}_t i_{t+j}\right\}, \tag{5}$$

where $\mu_{1,2} = \frac{1}{2}\left[1 + \beta + 4\kappa\sigma \pm \sqrt{(1+\beta+4\kappa\sigma)^2 - 4\beta}\right]$, with $\mu_1 > 1 > \mu_2 > 0$, and C is an arbitrary constant that indexes the model's multiple equilibria. The continuous-time solution is given by

$$\pi(t) = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2}\left\{Ce^{-\lambda_2 t} + \int_0^t e^{-\lambda_2 s} i(t-s)\,ds + \int_0^\infty e^{-\lambda_1 s} i(t+s)\,ds\right\}, \tag{6}$$

where $\lambda_{1,2} = \frac{1}{2}(\pm\rho + \sqrt{\rho^2 + 4\kappa\sigma})$, with $\lambda_1 \geq \lambda_2 > 0$, and C is, as above, a constant of integration that indicates the multiplicity of equilibria for any given path for nominal

interest rates.

The continuous-time solution makes it clear how indeterminacy manifests itself mathematically. The full solution to second-order differential equation (or, equivalently in this case, a system of first-order differential equations) requires two boundary conditions. Standard practice imposes a return to steady-state, which implies a boundary condition at infinity; indeterminacy arises because nothing else (yet) in the model determines the initial condition for inflation or the output gap.[9] The no-inflation-jump equilibrium, for example, assumes that $\pi(0) = 0$. Simply plug in $t = 0$ on the right hand side of equation 2.1 and set it equal to zero to obtain that $C = \int_0^\infty e^{-\lambda_1 s} i(s) ds$.

The model impulse responses in the last column of Figure 1 are constructed in a straightforward way from equation 2.1. With a known sequence of interest rates and the estimated (or assumed) initial value of inflation, the expression can be solved at each horizon of the impulse response.[10] To produce error bands, simply repeat the process for each bootstrap response of the interest rate.

The exercise described above is very simple; it only takes one paragraph to describe. But the reason it likely has not been examined in the simple model so far is because the equilibrium selection mechanism is not the standard one, which does not allow for arbitrary interest rate paths like those estimated from monetary policy shocks. (Although, as discussed below it is arguably more reasonable than the standard one.)

2.2 The standard equilibrium

The process described above is not typically how New Keynesian modelers pin down a unique equilibrium. Instead, they most often assume the central bank follows a monetary policy rule that obeys the Taylor principle like $i_t = \phi \pi_t + u_t$, with $\phi > 1$ and where u_t captures exogenous factors that might also influence the nominal rate. As described in King (2000) and later emphasized by Cochrane (2017) and subsequent papers, a unique equilibrium is achieved under this assumption because all equilibria *except one* lead to hyperinflation or hyperdeflation, and these are ruled out by assumption.

[9] It need not necessarily be an initial value or even a value at a specific time—it is equivalent to specifying any appropriate boundary condition when solving a differential equation. In section 2.2, below, I show how the Taylor principle implies a specific value for C in the standard New Keynesian equilibrium; in section 3 and 5, I use fiscal conditions to pin down C.

[10] All the responses in Figure 1 are estimated iteratively—either in a VAR or Romer and Romer's (2004) single-equation distributed lag model—so the impulse response can be computed to any arbitrarily long horizon, which is required to get an accurate estimate of the infinite sum, which necessarily must be truncated.

For example, in the continuous-time model, the monetary policy rule disturbance is typically assumed to follow a continuous-time AR(1): e.g., $u(t) = e^{-\eta t}$. The standard solution for inflation inherits the AR(1) properties of the only state variable $u(t)$, and so inflation has the form $\pi(t) = Au(t) = Ae^{-\eta t}$ for some constant A. The monetary policy rule then implies that $i(t) = (\phi A + 1)e^{-\eta t}$. As shown in Cochrane (2023), $\pi(t)$ will be negative and $i(t)$ will be positive as long as η is sufficiently large—that is, the shock $u(t)$ is not too persistent.

In this standard equilibrium, plugging the expressions above for $\pi(t)$ and $i(t)$ into equation 2.1 yields a unique value of C corresponding to this equilibrium. So, while the approach to finding a solution is somewhat different from the standard one, it nests the standard equilibrium as a particular case.

3 A model-based decomposition

In this section, I demonstrate how impulse responses in the New Keynesian model can be expressed as a weighted average of responses from two different simpler models that differ from each other in how agents form expectations within in the model. I proceed in continuous time for a few reasons. First, the continuous-time New Keynesian model is, essentially, a second-order non-homogeneous ordinary differential equation, the solutions of which have been well known for centuries; so, continuous time allows for simple closed-form solutions below. Second, in a few cases considered below and in the appendix, some properties of the Laplace transform and results from complex analysis are required for proofs; finally, and more prosaically, the model presented in Section 3.1 is the continuous-time limit of many such adaptive-expectations models that in discrete time have some differences in timing assumptions that disappear in the continuous-time limit.

The New Keynesian model in continuous time, as stated above, consists of an IS equation, resulting from a household consumption Euler equation and market clearing, and a Phillips curve resulting from profit-maximization by monopolistically competitive firms that face pricing frictions:

$$\begin{aligned}\mathbb{E}[\dot{x}] &= \sigma(i - \pi),\\ \mathbb{E}[\dot{\pi}] &= \rho\pi - \kappa x,\end{aligned}$$

where π is inflation, x is the output gap, etc. and I have used the short hand $\mathbb{E}[\dot{x}] \equiv \frac{1}{dt}\mathbb{E}[dx]$

and dropped time arguments for convenience.[11]

As shown in Cochrane (2018) and Cochrane (2023), for any path of nominal interest rates, the solutions to this model can be expressed as

$$\pi(t) = \Lambda \left\{ Ce^{-\lambda_2 t} + \lambda_2 \int_0^t e^{-\lambda_2 s} i(t-s)\, ds \right\} + (1-\Lambda) \left\{ \lambda_1 \int_0^\infty e^{-\lambda_1 s} i(t+s)\, ds \right\},$$

where $\lambda_{1,2} = \frac{1}{2}(\pm\rho + \sqrt{\rho^2 + 4\kappa\sigma})$, with $\lambda_1 \geq \lambda_2 > 0$, $\Lambda = \lambda_1/(\lambda_1 + \lambda_2)$, and C is an as-yet undetermined coefficient. This coefficient captures mathematically the well-known indeterminacy issue in New Keynesian models; each value of C reflects a different equilibrium response of inflation to the path of interest rates that is consistent with the model equations.

The model evidently features both forward- and backward integrals: inflation at time t is a discounted moving average of past and future nominal interest rates. It is tempting to say that the forward integral reflects forward-looking behavior by agents in the model, while the backward integral captures the lingering effects of past interest rates ("backward-looking" behavior). But these are convenient mathematical expressions of the *solution* to the model, and do not necessarily correspond directly to anything about agents' behavior within the model. We are interested in how expectations drive inflation dynamics, and it is not at all obvious what each of these terms in the models' solution represents.[12]

To get at that, as I show below, the full New Keynesian model can be expressed as an "average" of two different models: one with agents with rational expectations and one with adaptive expectations. We will see, somewhat counterintuitively, that the first integral in the expression above involving past interest rates actually captures the dynamics of a forward-looking, rational expectations model, while the last term involving future interest rates captures the dynamics of a backward-looking, adaptive expectations model. An integral of future values is the solution to a completely backward-looking model because of the requirement of stability. How the first term, the decaying exponential with the undetermined coefficient, is attributed to each model depends on the interest rate path being studied.

[11] These can be derived informally simply by taking the limit as the length of the time period Δt goes to zero in the discrete time model

$$\begin{aligned} x_t &= \mathbb{E}_t x_{t+\Delta t} - \sigma(i_t - \mathbb{E}_t \pi_{t+\Delta t})\Delta t \\ \pi_t &= \beta \mathbb{E}_t \pi_{t+\Delta t} + \kappa x_t \Delta t, \end{aligned}$$

or formally from continuous-time primitives.

[12] Blanchard (1979) discusses the distinction between forward and backward solutions and forward- and backward-loooking behavior within a model. McCallum (2011) discusses the connections between solutions of and expectations and causality within linear rational expectations models more generally.

To see this, let's first describe two slightly simpler models.

3.1 An "Old" Keynesian adaptive-expectations model

First, consider the model from Cochrane (2024) that consists of a static IS curve and backward-looking Phillips curve:

$$x = -\sigma(i - \pi)$$
$$\dot{\pi} = \kappa x.$$

Note that the Phillips curve now says that the actual derivative (not its expectation) is proportional to the output gap. This reflects adaptive expectations: past values of inflation determine inflation today and so inflation becomes a state variable.[13] These two expressions are the continuous time limit of many simple adaptive-expectations models of monetary policy.[14] The Philips curve also says inflation increases when the output gap is positive.[15] One could include a $\rho\pi$ term on the right-hand side, but it complicates the algebra without illuminating much about the decomposition below. For simplicity in what follows, let $\alpha \equiv \kappa\sigma$.

As is well known, this model is unstable. A generic path of interest rates will send inflation spiraling off to $\pm\infty$. This is reflected when you solve the model backward to get that $\pi(t) = \alpha \int_0^t e^{\alpha s} i(t-s) ds$. While this solution will be correct for any generic interest rate, any stable, non-hyperinflationary solution will also satisfy the forward solution plus an initial condition $\pi(t) = \alpha \int_0^\infty e^{-\alpha s} i(t+s) ds$ and $\pi(0) = 0$. The initial conditions that $\pi = 0$ arises because under adaptive expectations inflation is a state variable that cannot "jump" after an unanticipated shock. Therefore, we conclude that

$$\pi_{\text{AE}}(t) = \alpha \int_0^\infty e^{-\alpha s}\, i(t+s)\, ds \tag{7}$$

captures inflation dynamics in the "Old" Keynesian adaptive-expectations model.[16]

[13] We would obtain this Philips curve as $\Delta t \to 0$ in $\pi_t = \pi_{t-\Delta t} + \kappa x_t \Delta t$, a standard backward-looking Philips curve.

[14] These include Hall and Mankiw (1994), Taylor (1994, 1999), and Hetzel (2000), as well as the textbook intermediate macroeconomic models described in, e.g., Jones (2021) and Mankiw (2022).

[15] Although this may reflect how policymakers view the Philips curve, this model is not microfounded like the New Keynesian model is.

[16] In the appendix, I show that the standard way of obtaining stability in this model—including a Taylor rule $i_t = \phi\pi_t + u_t$ with $\phi > 1$—also satisfies these conditions. Equation 7 shows that other interest rate paths that lead to stability are also possible.

3.2 A simple rational expectations model

Next consider a simplified version of the New Keynesian model, also from Cochrane (2024), specifically one with the same static IS curve as in the adaptive expectations model, and with $\rho = 0$ in the New Keynesian Phillips curve:

$$\begin{aligned} x &= -\sigma(i - \pi) \\ \mathbb{E}[\dot{\pi}] &= -\kappa x, \end{aligned}$$

Unlike the adaptive-expectations model, this simple rational-expectations model is stable for any interest rate path, but like the full New Keynesian model, it is indeterminate. Its solutions are indexed by a constant of integration $\pi(t) = Ce^{-\alpha t} + \int_0^t e^{-\alpha s} i(t-s) ds$. In this simple the model $C = \pi(0)$, but that is not the case in the full model. Therefore, we conclude that

$$\pi_{\text{RE}}(t) = Ce^{-\alpha t} + \int_0^t e^{-\alpha s}\, i(t-s)\, ds. \tag{8}$$

captures inflation dynamics in the simple New Keynesian rational-expectations model.

3.3 A full New Keynesian model with $\rho = 0$

If we continue, as in the previous subsections, to assume that $\rho = 0$—again, just to simplify the algebra a bit—then plugging into the solution to the full model above gives

$$\pi_{\text{NK}}(t) = \frac{1}{2}\left\{ \underbrace{Ce^{-\sqrt{\alpha} t} + \sqrt{\alpha} \int_0^t e^{-\sqrt{\alpha} s}\, i(t-s)\, ds}_{\text{Rational Expectations}} \right\} + \frac{1}{2} \underbrace{\sqrt{\alpha} \int_0^\infty e^{-\sqrt{\alpha} s}\, i(t+s)\, ds}_{\text{Adaptive Expectations}}, \tag{9}$$

and we see that inflation dynamics in the full model can be decomposed into an average of two separate pieces: one coming from a simple rational expectations model with a static Phillips curve and another coming from the same simple model under adaptive expectations. This attribution of the two pieces of the New Keynesian model to the two submodels makes internal sense for any interest rate path that is admissible in both models. But a generic interest rate path leads to inflation spiraling off to infinity in the adaptive expectations model. The question remains then, to what models do we attribute these various pieces of the New Keynesian model for a generic interest rate. I discuss this issue below.

	Three Models		
Equations	Adaptive Expectations (**AE**)	Rational Expectations (**RE**)	New Keynesian (**NK**)
IS Curve	$x = -\sigma(i - \pi)$	$x = -\sigma(i - \pi)$	$\mathbb{E}[\dot{x}] = \sigma(i - \pi)$
Phillips Curve	$\dot{\pi} = \kappa x$	$\mathbb{E}[\dot{\pi}] = -\kappa x$	$\mathbb{E}[\dot{\pi}] = -\kappa x$
Solution: $\pi(t) =$	$\alpha \int_0^\infty e^{-\alpha s} i_{t+s} ds$	$\pi(0)e^{-\alpha t} + \alpha \int_0^t e^{-\alpha s} i_{t-s} ds$	$\frac{1}{2}\{Ce^{-\sqrt{\alpha}t} + \sqrt{\alpha}\int_0^t e^{-\sqrt{\alpha}s} i_{t-s} ds$ $+\sqrt{\alpha}\int_0^\infty e^{-\sqrt{\alpha}s} i_{t+s} ds\}$

TABLE 1: Three models in continuous time. I have assumed $\rho = 0$ and defined $\alpha \equiv \kappa\sigma$ to simplify the algebra and notation. $\mathbb{E}$ denotes rational expectations.

3.4 Decomposition of model responses

To illustrate how to apply the decomposition to the empirical impulse responses, let's first look at model-based responses. The models are summarized in Table 1. Since it's simply an illustration, let's make things even simpler and set $\alpha = 1$ and focus on the no-inflation jump equilibrium $\pi(0) = 0$, since that is the assumption in most of the empirical exercises below. There are also a few other good reasons to focus on this equilibrium, which I discuss below. In any case, I will relax these assumptions later on. When $\alpha = 1$, we clearly have

$$\pi_{\text{NK}} = \frac{1}{2}\pi_{\text{RE}} + \frac{1}{2}\pi_{\text{AE}}.$$

As mentioned above, this makes internal sense for any interest rate path that leads to stability in the AE model. But an arbitrary interest rate will generically lead to instability in the AE model, and the decomposition becomes less sensible. In the sections below, I first illustrate the decomposition using an interest rate that is stable in the AE model, and then show how to attribute the pieces of Equation 9 to the underlying simpler models for an arbitrary interest rate path, using an AR(1) as an example.

3.4.1 An interest rate that leads to stability

Typically, model-based impulse responses are computed from an AR(1) shock. While it's perfectly natural and possible to look at the AR(1) interest rate path here, it makes sense at first to consider a path of interest rates that leads to stability in all three models. An AR(1) interest rate $i(t) = i(0)e^{-\eta t}$ will lead to instability in the Old Keynesian adaptive expectations model.[17] To see this, recall that any stable interest rate path must have $\pi(t) \propto$

[17] Unless it is taken in response to offset another shock, which we are not considering here (yet).

$\int_0^\infty e^{-\alpha s} i(t+s)ds$ and $\pi(0) = 0$, which means that $\int_0^\infty e^{-\alpha s} i(s)ds = 0$. For an AR(1) interest rate path, $i(t)$ never changes sign, so it cannot integrate to zero.

So, instead let's decompose the full model responses to an interest rate path given by the difference between two exponential functions:

$$i(t) = e^{-\beta t} - \frac{\alpha + \gamma}{\alpha + \beta} e^{-\gamma t}.$$

It is straightforward to show that the discounted integral of this function at time 0 gives $\pi_{\text{AE}}(0) = 0$ and hence leads to stability in the adaptive expectations model. This function is an appealing starting point for four other reasons. First, a function of this form—the difference between two exponentials—is in multiple senses the "closest" thing to an AR(1) that leads to stability in the AE model model.Second, it is also an interest rate path that requires no fiscal support to obtain in equilibrium in the AE model.[18] Third, after a monetary policy shock in the AE model with a Taylor rule, the interest rate follows a process of this form. And finally, the empirical interest rate responses to monetary policy shocks in Figure 1 all display similar dynamics to this process: the interest rate is first positive before later turning negative.

So, again recalling that we are setting $\alpha = 1$ for simplicity we get in the AE model

$$\begin{aligned}
\pi_{\text{AE}}(t) &= \int_0^\infty e^{-s} \left\{ e^{-\beta(t+s)} - \frac{1+\gamma}{1+\beta} e^{-\gamma(t+s)} \right\} ds \\
&= \frac{1}{1+\beta} \left(e^{-\beta t} - e^{-\gamma t} \right).
\end{aligned}$$

In the simple RE model we get

$$\begin{aligned}
\pi_{\text{RE}}(t) &= \pi(0) e^{-t} + \int_0^t e^{-s} \left\{ e^{-\beta(t-s)} - \frac{1+\gamma}{1+\beta} e^{-\gamma(t-s)} \right\} ds \\
&= \frac{1}{1+\beta} \left\{ \frac{1+\beta}{1-\beta} e^{-\beta t} - \frac{1+\gamma}{1-\gamma} e^{-\gamma t} - \frac{\beta - \gamma}{(1-\beta)(1-\gamma)} e^{-t} \right\},
\end{aligned}$$

where the no-inflation-jump equilibrium imposes $\pi(0) = 0$. Therefore, the New Keynesian

[18] I discuss this more in Section ZQBD. In fact, this interest rate path together with the no-inflation-jump equilibrium selection also requires no fiscal support in either the RE or NK models.

model's inflation response is the simple average of these two simpler model responses:

$$\begin{aligned}\pi_{\text{NK}}(t) &= \frac{1}{2}\left\{\frac{1}{1+\beta}\left(e^{-\beta t}-e^{-\gamma t}\right)+\frac{e^{-\beta t}-e^{-t}}{1-\beta}-\frac{1+\gamma}{1+\beta}\frac{e^{-\gamma t}-e^{-t}}{1-\gamma}\right\}\\ &= \frac{1}{1+\beta}\left\{\frac{1}{1-\beta}e^{-\beta t}-\frac{1}{1-\gamma}e^{-\gamma t}-\frac{\beta-\gamma}{(1-\beta)(1-\gamma)}e^{-t}\right\},\end{aligned}$$

The three responses are displayed in Figure 2. It's clear that the full model response always lies exactly between the OK and RE models when $\rho = 0$ and $\alpha = 1$. It's also apparent that the RE model exhibits a price puzzle while the AE model does not, a preview of what we will find in the empirical decompositions: the model attributes the price puzzle almost entirely to rational expectations. Their average gives a flat initial inflation response in the full model to the interest rate shock.[19]

3.5 Which submodel accounts for most of the response

How should we quantify which piece of the NK model—the RE piece or the AE piece—accounts for the response? Some cases are pretty clear. For example in the no-fiscal-support equilibrium response to the AR(1) interest rate, the portion of the response coming from the RE model is almost equal to the NK response. A natural choice would be say what fraction of the NK response is made up by each submodel response, but that runs into issues with zeros in the denominator. So, I propose a slightly different metric. What fraction of the total distance from the NK model of the AE and RE models is accounted for by each? Specifically, define

$$\Delta_{\text{AE}}(t) \equiv |\pi_{\text{NK}}(t) - \pi_{\text{AE}}(t)| \qquad \text{and} \qquad \Delta_{\text{RE}}(t) \equiv |\pi_{\text{NK}}(t) - \pi_{\text{RE}}(t)|,$$

and then we interpret the expressions

$$\mathcal{F}_{\text{RE}}(t) \equiv 1 - \frac{\Delta_{\text{RE}}(t)}{\Delta_{\text{RE}}(t) + \Delta_{\text{AE}}(t)} \qquad \text{and} \qquad \widetilde{\mathcal{F}}_{\text{RE}} \equiv \frac{1}{T}\int_0^T \mathcal{F}_{\text{RE}}(t)\,dt$$

[19] Also worth noting is that the interest rate intersects the inflation responses at the peak of the RE model, the trough of the AE model and the inflection point of the NK model. Note that this arises because the AE model is $(D-\alpha)\pi(t) = -\alpha i(t)$, and setting $i(t) = \pi(t)$, we get that $D\pi(t) \equiv \pi'(t) = 0$; we get the same thing for the RE model with $(D+\alpha)\pi(t) = \alpha i(t)$ implying the same thing, whereas in the full NK model we have $(D^2+\alpha)\pi(t) = \alpha i(t)$ which implies that $D^2\pi(t) \equiv \pi''(t) = 0$. I am not sure yet of the economic significance of this result or if it's a mathematical curiosity.

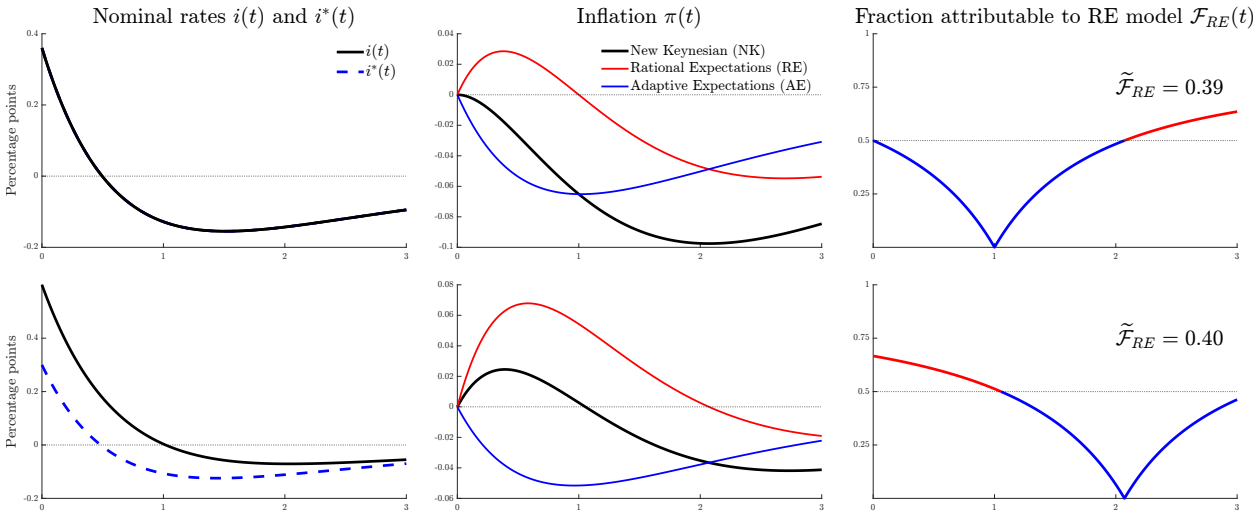


FIGURE 2: DECOMPOSITION OF THE NO-INFLATION-JUMP EQUILIBRIUM, $\rho = 0$ AND $\alpha = 1$

Notes: The responses of inflation to different interest rate shocks in the New Keynesian (NK), rational expectations (RE), and adaptive expectations (AE) models in equilibria in which inflation's response is zero on impact. First row, $i(t) = e^{-\beta t} - \frac{\alpha+\gamma}{\alpha+\beta}e^{-\gamma t}$; second row $i(t) = e^{-\beta t} - Ae^{-\gamma t}$ and $i^*(t)$ is the closest interest rate path to $i(t)$ that leads to stability in the AE model. Middle column illustrates how the NK model is the simple average of the two models; in the right column the NK response is the sum of AE and RE responses. Both rows depict in the "no-inflation-jump" equilibrium; the bottom row exhibits a price puzzle while the top row does not. Responses displayed for $\rho = 0$, $\alpha = 1$, $\beta = 1.5$, $\gamma = 0.6$, and $A = 0.5$.

as the fraction of the NK response accounted for by the RE model at time t and on average over the full response, where T is the horizon of the response. The same values for the AE model are obtained by subtracting those above from one. These expressions have intuitive appeal: $\mathcal{F}_j(t) = 1$ means model j is equal to NK at time t, while a value of zero means the other submodel is, and a value of 0.5 means that each submodel response is the same distance from the NK response. Moreover $\widetilde{\mathcal{F}}_j = 1$ means the model-j response is the same as the NK response throughout the entire response horizon.

3.5.1 An interest rate that leads to instability

Now consider an arbitrary interest rate path $i(t)$. This will generically lead to instability in the AE model, so it makes little sense to attribute the forward integral in Equation 9 to inflation dynamics from that model since it goes off to infinity. There are two natural ways to address this issue, each of which is also feasible to apply in empirical settings. One is simply to attribute a piece of the exponentially decaying term $Ce^{-\sqrt{\alpha}t}$ to the forward integral so that the implied $\pi_{\text{AE}}(0) = 0$. The sum of these terms can then be interpreted as the response of the AE model to a slightly different interest rate $\widehat{i}(t)$, one that can be solved

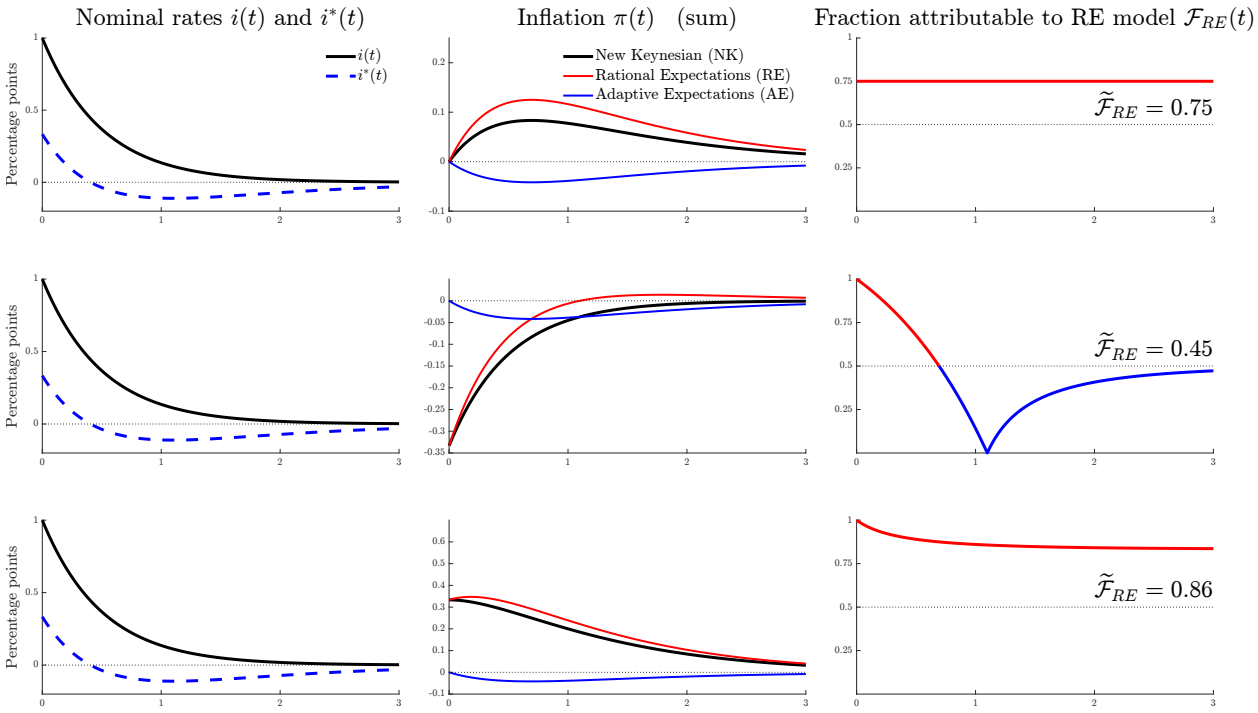


FIGURE 3: DECOMPOSITION OF INFLATION'S RESPONSE, $\rho = 0$ AND $\alpha = 1$

Notes: The response of inflation in the New Keynesian (NK) model to the interest rate $i(t) = e^{-\beta t} - \frac{\alpha+\gamma}{\alpha+\beta}e^{-\gamma t}$ is the simple average of the responses of the simple model with rational expectations (RE) and adaptive expectations (AE) when $\alpha = 1$ and $\rho = 0$. Responses displayed for $\beta = 1.5$ and $\gamma = 0.6$.

for by combining Equations 7 and 9.

Another approach is to break down the inflation response to an arbitary interest rate path $i(t)$ of the full model into the response to the same interest rate path in the simple RE model, and to $i^*(t)$ in the AE model, where $i^*(t)$ is the interest rate path that is *closest* to $i(t)$, in a sum-of-squared-errors sense, that leads to a stable solution. Specifically, given any arbitrary interest rate path $i(t)$, we want to find $i^*(t)$ to solve the following problem:

$$\min_{i^*(t)} \frac{1}{2} \int_0^\infty e^{-\delta t} \left(i^*(t) - i(t)\right)^2 dt$$

$$\text{subject to } \int_0^\infty e^{-\alpha t} i^*(t)\, dt = 0,$$

where $\delta \geq 0$ is included in case one wishes to discount interest rate differences further in the future.

I show each of these approaches using an AR(1) interest rate path $i(t) = i(0)e^{-\eta t}$. The

solution of the NK model for this interest rate is

$$\pi_{\text{NK}}(t) = \frac{1}{2}\left\{Ce^{-t} + i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta} + i(0)\frac{1}{1+\eta}e^{-\eta t}\right\}, \tag{10}$$

where the second and third terms in the sum are the solved backward and forward integrals from Equation 9. Clearly, the third term evaluated at $t = 0$ is positive, so it does not correspond to a stable solution to the AE model. To begin, consider first the no-inflation-jump equilibrium with $\pi_{\text{NK}}(0) = 0$, which implies that $C = -i(0)/(1+\eta)$. I focus on this case first since most of the empirical responses in Figure 1 assume it, but I consider two different equilibria below.

In the first approach, we "shift" a portion of the decaying exponential Ce^{-t} to the forward integral $i(0)/(1+\eta)e^{-\eta t}$ such that the sum of the two terms at $t = 0$ is zero. In this simple example, it turns out that *all* of the the exponential must be shifted to the forward integral,[20] but that will not generally be the case as we will see in the next example. So, now we decompose the expression above as follows

$$\pi_{\text{NK}}(t) = \frac{1}{2}\Bigg\{ \underbrace{i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta}}_{\substack{\equiv \pi_{\text{RE}}(t) \\ \text{Rational Expectations}}} + \underbrace{i(0)\frac{e^{-\eta t} - e^{-t}}{1+\eta}}_{\substack{\equiv \widehat{\pi}_{\text{AE}}(t) \\ \text{Adaptive Expectations}}} \Bigg\}, \tag{11}$$

where $\pi_{\text{RE}}(t)$ is the no-inflation jump equilibrium response to $i(t)$ in the RE model and $\widehat{\pi}_{AE}(t)$ is the response of the AE model to $\widehat{i}(t)$, which can be solved for using Equation 7. In this example, $\widehat{i}(t) = i(0)\left(e^{-\eta t} - \frac{2}{1+\eta}e^{-t}\right)$, which, it is worth noting, is of the same form as the interest rate path in Section 3.4.1, the difference between two exponentials with $\beta = \eta$ and $\gamma = 1$. The decomposition based on this method of "shifting" the exponential decay term is displayed in the second panel of Figure 2.

Now, the second sensible approach in which we find $i^*(t)$ *closest* to $i(t)$, is straightforward to solve using Lagrange multipliers, which I show in the Appendix, leads to the following solution when $i(t) = i(0)e^{-\eta t}$:

$$i^*(t) = i(0)\left(e^{-\eta t} - \frac{2\alpha - \delta}{\alpha + \eta}e^{(\delta-\alpha)t}\right).$$

Since we have assumed $\alpha = 1$, and if we also set $\delta = 0$, we get that $i^*(t) = \widehat{i}(t)$, so we get

[20] To see this, note that $-i(0)/(1+\eta)e^{-t} + i(0)/(1+\eta)e^{-\eta t} = 0$ at $t = 0$.

the same answer as in Equation 11; that is

$$\pi_{\text{NK}}(t) = \frac{1}{2}\pi_{\text{RE}}(t) + \frac{1}{2}\widehat{\pi}_{\text{AE}}(t) = \frac{1}{2}\pi_{\text{RE}}(t) + \frac{1}{2}\pi^*_{\text{AE}}(t).$$

This decomposition does not depend on the choice of the no-inflation-jump equilibrium. To see this, consider instead the standard equilibrium in the NK model, in which a Taylor rule of the form $i = \phi\pi + u$ with $\phi > 1$ is added to the IS and Phillips curve. This leads to a unique non-explosive equilibrium; if, as is usually assumed, the disturbance on the Taylor rule u follows an AR(1) $u(t) = u(0)e^{-\eta t}$, then the nominal interest rate and inflation also are AR(1)s proportional u . Rather than rederiving the decomposition for this case, I simply select the equilibrium for $i(t) = i(0)e^{-\eta t}$ that implies $\pi_{\text{NK}}(t) \propto i(t)$. From Equation 10, it is evident that $C = i(0)/(1-\eta)$ gives the desired result, for in that case we have

$$\pi_{NK}(t) = \frac{1}{2}\left\{i(0)\frac{1}{1-\eta}e^{-t} + i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta} + i(0)\frac{1}{1+\eta}e^{-\eta t}\right\} = i(0)\frac{1}{1-\eta^2}e^{-\eta t} = \frac{1}{1-\eta^2}i(t),$$

where despite the simplification, we can still use the first equality to construct the decomposition. As before, we have two ways to allocate the first term in the triple sum (the decaying exponential) to the third term (the forward integral).

Using the first approach where we shift just enough of the first term to the third so that the implied $\widehat{\pi}_{\text{AE}}(0) = 0$, we get, as we did in the no-inflation-jump equilibrium, that the requisite portion is $-i(0)/(1+\eta)e^{-t}$. The difference is that in the standard equilibrium, a portion of the exponential decay will still be attributed to the RE model since $-i(0)/(1+\eta)$ is no longer equal to C. Specifically $i(0)[1/(1-\eta) + 1/(1+\eta)]e^{-t}$ of the decay term will be associated with the RE model. Therefore, we have the following expression illustrating the decomposition of the standard equilibrium of the NK model:

$$\pi_{\text{NK}}(t) = \frac{1}{2}\Bigg\{\underbrace{i(0)\frac{1}{1-\eta^2}e^{-t} + i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta}}_{\equiv\pi_{\text{RE}}(t)\text{ Rational Expectations}} + \underbrace{i(0)\frac{e^{-\eta t} - e^{-t}}{1+\eta}}_{\equiv\widehat{\pi}_{\text{AE}}(t)\text{ Adaptive Expectations}}\Bigg\}. \tag{12}$$

This decomposition is displayed in the third row of Figure 2. Note also that since $\widehat{\pi}_{\text{AE}}$ is the same, we will have the same values as in the no-inflation-jump equilibrium for $\widehat{i} = i^*$ and $\pi^*_{\text{AE}} = \widehat{\pi}_{\text{AE}}$.

At the risk of over-exampling, let's do one more decomposition. Monetary policy actions in the NK model require fiscal backing: the fiscal authority must adjust taxes and spending

to bring the present discounted value of future primary surpluses in line with real interest rate. Typically, and in all the equilibria constructed so far, researchers assume that fiscal policy is "passive," or more generously "locally Riccardian"[21]; that is, the fiscal authority adjusts policy to realize the central bank's equilibrium selection irrespective of any potential goals of fiscal policy.

A natural alternative, one pursued in the fiscal theory of the price level (see Cochrane (2023) and Cochrane (2024)), is to study the equilibrium in which there is no change in the present discounted value of primary surpluses for a given path of interest rates. This approach recognizes that in equilibrium, in the nonlinear version of the model, real value of outstanding nominal government debt $B(t)$ is given by

$$V(t) \equiv \frac{B(t)}{P(t)} = \int_0^\infty \frac{\lambda(t+\tau)}{\lambda(t)}\, s(t+\tau)\, d\tau,$$

where $P(t)$ is the price level, $\lambda(t)$ is the household stochastic discount factor, and $s(t)$ is the government's real primary budget surplus. After linearization, this expression becomes

$$\dot{v} = \rho v + i - \pi - \widetilde{s}, \tag{13}$$

where $v(t)$ and $\widetilde{s}(t)$ are (deviations from steady state of) the real value of outstanding government debt and real primary surplus scaled by the steady-state value of the former. In the model studied so far, there is only short-term debt, so like inflation in the AE mode, v is a state variable that cannot "jump" on the impact of a shock and hence $\dot{v}$ rather than $\mathbb{E}[\dot{v}]$ appears in Equation 13. Integrating the expression forward gives

$$v(t) = \int_0^\infty e^{-\rho(t+\tau)}\, \left[\widetilde{s}(t+\tau) - (i(t+\tau) - \pi(t+\tau))\right]\, d\tau,$$

which together with the fact that v cannot jump on impact, implies that

$$\int_0^\infty e^{-\rho t}\, (i(t) - \pi(t))\, dt = \int_0^\infty e^{-\rho t}\, \widetilde{s}(t)\, dt \tag{14}$$

If fiscal policymakers do not adjust surpluses after a change in nominal interest rates, i.e., $\widetilde{s}(t)$ is zero for all t,[22] then Equations 9 and 14 pin down a unqiue equilibrium. Equation 9 is plugged into Equation 14 to find the unique undetermined coefficient C that satisfies both equations.

[21]The former is Leeper's (1991) term; the latter is Woodford's (2003).

[22]Surpluses being unchanged is not required; any known function $\widetilde{s}(t)$ will work just as easily

As I hinted at above, when the interest rate is given by $i(t) = e^{-\beta t} - \frac{\alpha+\gamma}{\alpha+\beta}e^{-\gamma t}$, which is stable in the AE model, the no-inflation-jump equilibrium is the outcome that would obtain if instead the equilibrium selected was based on a no-fiscal-support condition (i.e., $\widetilde{s}(t) = 0 \ \forall t$).[23] But when the interest rate is an AR(1), it will not be the case that the no-inflation-jump equilibrium will correspond to the no-fiscal-support equilibrium.

Starting from the general solution to the NK model when $i(t) = i(0)e^{-\eta t}$, Equation 10 the no-fiscal-support condition requires

$$\frac{1}{2}\int_0^\infty \left\{ Ce^{-t} + i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta} + i(0)\frac{1}{1+\eta}e^{-\eta t} \right\} dt = i(0)\int_0^\infty e^{-\eta t}\, dt$$
$$\frac{1}{2}\left\{ C + i(0)\frac{1}{\eta(1-\eta)} - i(0)\frac{1}{1-\eta} + i(0)\frac{1}{\eta(1+\eta)} \right\} = \frac{i(0)}{\eta},$$

which gives $C = i(0)/(1+\eta)$ and hence

$$\pi_{\text{NK}}(t) = \frac{1}{2}\left\{ i(0)\frac{1}{1+\eta}e^{-t} + i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta} + i(0)\frac{1}{1+\eta}e^{-\eta t} \right\} = \frac{i(0)}{1-\eta^2}(e^{-\eta t} - \eta e^{-t}).$$

We again must either allocate a portion of the decaying exponential to the AE model or find the response of the AE model to the closest interest rate to an AR(1) that leads to a stable solution, but as shown above, these will give the same answer. As above, we need to allocate $-i(0)/(1+\eta)e^{-t}$ to the AE portion and therefore leave $2i(0)/(1+\eta)$ with the RE model. Then we have

$$\pi_{\text{NK}}(t) = \frac{1}{2}\Bigg\{ \underbrace{2i(0)\frac{1}{1+\eta}e^{-t} + i(0)\frac{e^{-\eta t} - e^{-t}}{1-\eta}}_{\equiv \pi_{\text{RE}}(t)\text{ Rational Expectations}} + \underbrace{i(0)\frac{1}{1+\eta}e^{-\eta t} - i(0)\frac{1}{1+\eta}e^{-t}}_{\widehat{\pi}_{\text{AE}}(t)\text{ Adaptive Expectations}} \Bigg\}. \tag{15}$$

This decomposition is displayed in the last row of Figure 2.

To sum up, in this section, we have seen that regardless of the method of equilibrium selection—be it a Taylor rule, a fiscal constraint, or an arbitrary choice by the researcher—the NK model can always be expressed the average of two models with static IS curves, one with adaptive expectations and one with rational expectations. Throughout this section, I have assumed $\alpha \equiv \kappa\sigma = 1$ and $\rho = 0$. In the next section, I show how the decomposition can be generalized to alternative parameterizations. Different parameterizations do not present

[23] It is straightforward to verify that for this interest rate $\int_0^\infty i(t)dt = \int_0^\infty \pi_{\text{NK}}(t)dt$ in the no-inflation-jump equilibrium.

any inherent difficulties, simply a bit more math and a slightly different interpretation.

3.6 Model decompositions when $\alpha \neq 1$ and $\rho > 0$

When we depart from the special case when $\alpha = 1$ and $\rho = 0$, the inflation repsonse of NK model is no longer the simple average of the AE and RE models; instead, it is a weighted average of AE and RE models in which the timing of interest rate changes has been shifted. To see this, let's first continue to assume that $\rho = 0$, but let α be any positive number. In this case, inflation is given by Equation 9, repeated here for convenience:

$$\pi_{\text{NK}}(t) = \frac{1}{2}\left\{\underbrace{Ce^{-\sqrt{\alpha}t} + \sqrt{\alpha}\int_0^t e^{-\sqrt{\alpha}s}\, i(t-s)\, ds}_{\text{Rational Expectations}}\right\} + \frac{1}{2}\underbrace{\sqrt{\alpha}\int_0^\infty e^{-\sqrt{\alpha}s}\, i(t+s)\, ds}_{\text{Adaptive Expectations}}.$$

One way to interpret this expression when $\alpha \neq 1$ is that the NK model is the simple average of the RE and AE models with slightly different parameterizations, one with $\kappa_{\text{RE}}\sigma_{\text{RE}} = \kappa_{\text{AE}}\sigma_{\text{AE}} = \sqrt{\kappa_{\text{NK}}\sigma_{\text{NK}}}$. Another interpretation, which relies on some convenient properties of the Laplace transform is that the two pieces are the responses of the RE and AE models with the same calibration as the NK model to a *time-scaled* version of the same interest rate process. Applying a well-known property of the Laplace transform, one can using a change of the variable of integration show that

$$\sqrt{\alpha}\int_0^\infty e^{-\sqrt{\alpha}s}\, i(t+s)\, ds = \alpha\int_0^\infty e^{-\alpha s}\, i(t+\sqrt{\alpha}s)\, ds,$$

and

$$\sqrt{\alpha}\int_0^t e^{-\sqrt{\alpha}s}\, i(t-s)\, ds = \alpha\int_0^{\frac{t}{\sqrt{\alpha}}} e^{-\alpha s}\, i(t-\sqrt{\alpha}s)\, ds$$

so that the interest rate effects on inflation in the full model are felt either more or less quickly than, in this example, the AE model depending on whether α is greater or less than one.

(Express inflation in NK as average of these time-scaled model responses).

If $\rho > 0$, then inflation in NK becomes of a weighted average of the two different models. Use the fact that

$$\lambda_1\int_0^\infty e^{-\lambda_1 s}\, i(t+s)\, ds = \frac{\rho+\alpha}{\alpha}\underbrace{\alpha\int_0^\infty e^{-(\rho+\alpha)s}\, i\left(t+\frac{\rho+\alpha}{\lambda_1}s\right) ds}_{\text{time-scaled AE response}}$$

The full model decomposition can therefore be written as

$$\pi_{\text{NK}}(t) = \Lambda\frac{\alpha-\rho}{\alpha}\Big\{\underbrace{\widetilde{C}e^{-(\alpha-\rho)\widetilde{t}} + \alpha\int_0^{\widetilde{t}} e^{-(\alpha-\rho)s}\, i(t-\widetilde{s}_2)\, ds}_{\text{time-scaled RE model response}}\Big\}$$
$$+ (1-\Lambda)\frac{\alpha+\rho}{\alpha}\Big\{\underbrace{\alpha\int_0^{\infty} e^{-(\alpha+\rho)s}\, i(t+\widetilde{s}_1)\, ds}_{\text{time-scaled AE model response}}\Big\},$$

where $\widetilde{C} = \alpha/(\alpha-\rho)C$, $\widetilde{t} \equiv \lambda_2/(\alpha-\rho)t$, $\widetilde{s}_1 \equiv (\alpha+\rho)s/\lambda_1$, and $\widetilde{s}_2 \equiv (\alpha-\rho)s/\lambda_2$.[24] So, the NK model can, even in the case when $\alpha \neq 1$ and $\rho > 0$, be expressed as a weighted average of time-scaled responses in the simple RE and AE models. This time-scaling makes some intuitive sense: the more flexible are prices (i.e., the higher is α), the larger the effect on current inflation of further out interest rates.

Alternatively, we could interpret this time-scaling as simply giving us the interest rate in each submodel that produces each piece of the full model.

4 Empirical decompositions

Now it straightforward to apply the model-based decompositions to the empirical responses of nominal rates to monetary policy shocks. These are displayed in Figure 4.

The simple model-based response with a price puzzle looks qualitiatively similar to all the empirical responses, although it most closely resembles that from Christiano et al. (1999). In all cases the price puzzle periods are dominated by the effect coming from the simple ration expectation **RE** model.

5 The forward-looking Fisher effect

The forward-looking Fisher effect is the notion that higher expected *future* nominal interest rates can raise current inflation. While it is actually straightforward to see in the solutions above ($\pi(t)$ in equations 2.1 and 2.1 is increasing in $i_t + j$ for all j.), one can also see it in the standard two-equation set up.

[24] The presence of ρ in the terms in brackets, whereas it was not there before, arises because right-hand side of the Phillips curve in each model is $\rho\pi \pm \kappa x$ when $\rho \neq 0$. I assume that $\rho < \alpha$, which will hold unless prices are extremely sticky.

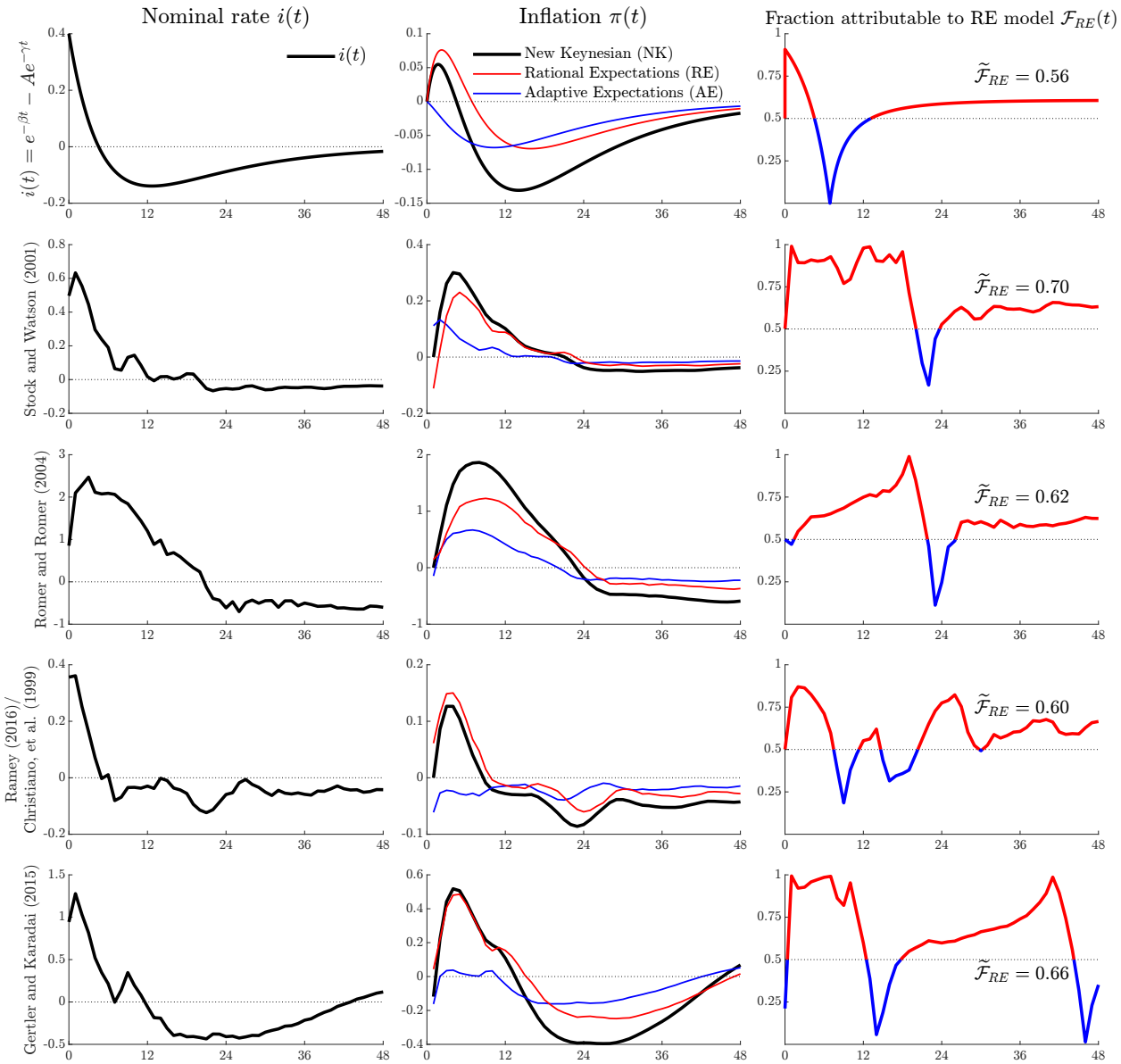


FIGURE 4: DECOMPOSITION OF MODEL AND EMPIRICAL RESPONSES

Notes: The response of inflation in the New Keynesian (NK) model to interest rate responses to monetary policy shocks. In the first row, $\beta = 0.1875$, $\gamma = 0.0375$, $A = 0.6$, while in all others the interest rate path is given the empirical point estimate. Parameters are the same as in Figure 1.

It is simplest to see the forward-looking Fisher effect in the discrete-time model under the simple parameterization that $\kappa = \sigma = 1$. This assumption simplifies the algebra a bit without losing much generality.[25] Consider the case when the central bank increases the

[25]Technically, $\beta = 1$ ensures long-run neutrality when nominal rate changes are permanent. This assumption is not crucial for the result; it just makes the math a bit simpler.

nominal rate permanently so that $i_t = 1$ for all t. Suppose that there exists some finite time period T such that the effect is expected to be Fisherian; that is, $\mathbb{E}_t \pi_{T+j+1} = \mathbb{E}_t i_{T+j} = 1$ for all $j \geq 0$. The IS and PC equations evaluated at time $T+1$ become

$$x_{T+1} = x_{T+2}$$

$$\pi_{T+1} = 1 - x_{T+1}.$$

Since the latter equals 1 by assumption, we must have $x_{T+1} = x_{T+2} = 0$. The same argument implies that $x_{T+j} = 0$ for all $j \geq 0$. Given that $x_{T+1} = 0$, evaluating the IS and PC equations at time T gives $\pi_T = 0$. Repeating this process all the way back to time zero implies that $\pi_t = 1$ and $x_t = 0$ for all t. So, in this simple example, if the effect of a permanent change in the nominal rate is ever expected to be exactly Fisherian, it is *super* Fisherian, with $i_t = \pi_t = 1$ for all t.

Now suppose instead of being *exactly* Fisherian at time T, people in the model expect that $\mathbb{E}_t \pi_{T+j+1} = \mathbb{E}_t i_{T+j} - \varepsilon_j = 1 - \varepsilon_j$, for all $j \geq 0$, where $\{\varepsilon_j\}$ is a monotonically decreasing sequence of positive numbers less than one with $\lim_{j \to 0} \varepsilon_j = 0$. In words, by time T, the effects of the permanent interest rate change is expected to be *almost* Fisherian and approaching neutral effects asymptotically. This scenario is *close* to the first one, but inflation is expected to be only slightly below the nominal rate. Now for any $j \geq 0$,

$$x_{T+j} = x_{T+j+1} - (1 - (1 - \varepsilon_j)) \;\Rightarrow\; \Delta x_{T+j+1} = \varepsilon_j > 0.$$

So the output gap is strictly increasing during the near-Fisherian period. Together with long-run neutrality this fact implies that $x_{T+j} < 0$. The Phillip curve then gives

$$\pi_{T+j} = 1 - \varepsilon_j + x_{T+j} < 1 - \varepsilon_j < 1,$$

and in particular $\pi_T = 1 - \varepsilon_0 + x_T$. Now, at time $t = T - 1$, the IS curve gives

$$x_{T-1} = x_T - (1 - \pi_{T+1}) = x_T - (1 - (1 - \varepsilon_1)) = x_T - \varepsilon_1,$$

while the Phillips curve gives

$$\pi_{T-1} = \underbrace{1 - \varepsilon_0 + x_T}_{\pi_T} + \underbrace{x_T - \varepsilon_1}_{x_{T-1}} = 1 + 2x_T - (\varepsilon_0 + \varepsilon_1).$$

Consequently, $\pi_{T-1} < \pi_T$ (recall that $x_T < 0$) and it is straightforward to continue this process to show that inflation is lower in every time period in this scenario than the full Fisherian outcome considered above.

Thus, expectations of *lower* interest rates in the future, even in an arbitrarily small neighborhood of a fully Fisherian equilibrium, lower inflation now.[26]

The expressions used thus far are *equilibrium* conditions not behavioral ones. They are not demand equations, but supply-equals-demand equations. They do not describe how the behavior of people within the model bring about this equilibrium. In the simple cases considered above with a permanent increase in the interest rate, the first case, households expect no change in the real interest rate, and so they desire no change in consumption or in their labor supply. With no change in either labor supply or consumption demand, firms must produce the same amount with the same labor input, which implies no need to change the real price of consumption. This in turn implies that dollar price of consumption must grow at the same rate as the nominal rate.

In the case of a price-puzzle equilibrium, while higher real rates lower consumer demand on impact, firms (correctly) forecast a boom in output resulting from desired consumption *growth* when real rates are high. If prices were flexible, firms would lower their prices when the shock hits, and wait to raise them until the boom materialized; however, when prices are costly to adjust, keeping prices fixed initially and raising them in anticipation of the boom yields higher profits.

The previous two paragraphs attempt to describe the causal chain of what leads to these outcomes in equilibrium, but in a general equilibrium model everything is determined simultaneously. Moreover, because the New Keynesian model features multiple equilibria, all can be viewed as self fulfilling, at least in a sense. The expectations of a future boom cause prices not to fall in price-puzzle equilibria, and actions based on those expectations cause the expected outcome to actually happen. What caused the expectations to change in the first place need not be specified. In any case, the price puzzle equilibria displayed in Figure 1 make perfect sense to the people living in those model economies. The price puzzle is not a puzzle to them, but instead to macroeconomists and monetary policymakers.

The paths of inflation described in Equations 2.1 and 2.1 evidently features both forward-

[26]Readers may notice something like Del Negro, Giannoni and Patterson's (2023) "forward-guidance puzzle"—in this case, changes in inflation expectations further in the future have larger effects on inflation in the present. However, the forward-looking Fisher effect still exists (and indeed is even easier to demonstrate) in an extreme version of McKay, Nakamura and Steinsson's (2017) discounted Euler equation model in which there is no forward-guidance puzzle; this model is discussed in section 3, below.

and backward sums or integrals. But these are mathematical expressions of the *solution* to the model. We are interested in how expectations of agents within the model drive inflation dynamics, and it is not at all obvious what aspect each of these terms represent. To get at that, as I show below, the full New Keynesian model can be expressed as an "average" of two different models: one forward-looking with rational expectations and one backward-looking with adaptive expectations.

6 Endogenous responses to other shocks

Do the results above mean that central banks should *lower* interest rates to bring down inflation? Not necessarily. In this section I show that although a lower nominal rate will lower inflation if no other shock is present, in response to another shock that has increased inflation a higher rate can actually lower inflation, at least in the short run.

Now, as discussed above, in response to any shock, the central bank in the New Keynesian model can select any initial value of inflation it would like (or any initial value of output, or any value of output or inflation at any point along the impulse response, or any combination of the two). How then can we compare apples to apples? Although there are many ways, here I consider just one for illustrative purposes.

Suppose that a shock to the Phillips curve raises inflation by the amount of the shock on impact if the nominal interest rate does not change; but if instead the nominal rate *rises* in response to the shock, inflation rises by less than the size of the shock on impact. This criterion is a loose one, to be sure (by how much less does inflation rise?), but not totally unreasonable and at the very least consistent with the conventional intuition.

In this case, as long as the increase in nominal rates is neither too large nor too persistent, higher nominal rates offset the effects of the shock in the short run. That is, in spite of the results above, *in order to offset inflationary non-monetary shocks, central banks may indeed need to (endogenously) raise nominal interest rates.*

7 Discussion

One important omission depicted in not one figure in this paper is output. The output dynamics are totally at odds with the data. The same monetary policy shocks produce output responds that themselves reflect a puzzle. After a contractionary monetary policy shock, output rises and unemployment falls in the very short run, rather than jumping down

immediately and continuing to grow as implied by a standard consumption Euler equation. The subtleties of the output effects of monetary policy are beyond the scope of this paper, so I leave them for future research[27]

8 Conclusion

To the pretend people living in the New Keynesian model, the increase in prices following higher nominal interest rates is not confusing at all. But the price puzzle is indeed hard to understand for economists armed only with the conventional macroeconomic intuition. In this paper, I have shown how this conventional intuition diverges from that of the New Keynesian model that undergirds most frameworks for understanding the practice of monetary policy in modern economies. If one accepts that the conventional wisdom is accurately enough captured by the intermediate macroeconomics adaptive expectations model, then the difference between the that model and the NK model is that same model under rational expectations. I show that rational expectations plus the Fisher relation among real rates, nominal rates, and inflation lead to a *forward-looking* Fisher effect whereby future nominal rates affect inflation today.

The simple two-equation model does arguably a better job of matching the shape of prices' response to monetary shocks that larger models with more than ten times the number of equations, the intuition of which is opaque at best. This is a heretofore unappreciated strength of the simple sticky-price model that suggests a rethinking of the application of those larger models. Obviously, more research is needed to assess the speculations of this paragraph.

[27]Perhaps a paper titled, "The New Keynesian Quantity Puzzle: Reinterpreting Output Dynamics" is in order.

References

Blanchard, O. J. (1979): "Backward and Forward Solutions for Economies with Rational Expectations," *The American Economic Review*, 69, 114–118.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Chapter 2 Monetary policy shocks: What have we learned and to what end?" Volume 1 of Handbook of Macroeconomics: Elsevier, 65–148, https://doi.org/10.1016/S1574-0048(99)01005-8.

Cochrane, J. H. (2017): "The new-Keynesian liquidity trap," *Journal of Monetary Economics*, 92, 47–63, https://doi.org/10.1016/j.jmoneco.2017.09.003.

——— (2018): "Michelson-Morley, Fisher, and Occam: The Radical Implications of Stable Quiet Inflation at the Zero Bound," *NBER Macroeconomics Annual*, 32, 113–226, 10.1086/696050.

——— (2023): *The Fiscal Theory of the Price Level*, Princeton, New Jersey: Princeton University Press.

——— (2024): "Expectations and the neutrality of interest rates," *Review of Economic Dynamics*, 53, 194–223, https://doi.org/10.1016/j.red.2024.04.004.

Del Negro, M., M. P. Giannoni, and C. Patterson (2023): "The Forward Guidance Puzzle," *Journal of Political Economy Macroeconomics*, 1, 43–79, 10.1086/724214.

Eichenbaum, M. (1992): "'Interpreting the macroeconomic time series facts: The effects of monetary policy': by Christopher Sims," *European Economic Review*, 36, 1001–1011, `https://EconPapers.repec.org/RePEc:eee:eecrev:v:36:y:1992:i:5:p:1001-1011`.

Fry, R., and A. Pagan (2011): "Sign Restrictions in Structural Vector Autoregressions: A Critical Review," *Journal of Economic Literature*, 49, 938–60, 10.1257/jel.49.4.938.

Galí, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition* in , Economics Books (10495): Princeton University Press, `https://ideas.repec.org/b/pup/pbooks/10495.html`.

Gertler, M., and P. Karadi (2015): "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, 7, 44–76, 10.1257/mac.20130329.

Hall, R. E., and N. G. Mankiw (1994): "Nominal Income Targeting," in *Monetary Policy*: University of Chicago Press, 71–94.

Hetzel, R. L. (2000): "The Taylor Rule: Is It a Useful Guide to Understanding Monetary Policy," *Economic Quarterly*, 86, 1–33.

Jones, C. I. (2021): *Macroeconomics*, New York: W.W. Norton and Company, Inc. 5th edition.

King, R. (2000): "The new IS-LM model: language, logic, and limits," *Economic Quarterly*, 45–103, `https://EconPapers.repec.org/RePEc:fip:fedreq:y:2000:i:sum:p:45-103`.

Leeper, E. M. (1991): "Equilibria under active and passive monetary and fiscal policies," *Journal of Monetary Economics*, 27, 129–147, https://doi.org/10.1016/0304-3932(91)90007-B.

Mankiw, N. G. (2022): *Macroeconomics*, New York: Worth Publishers, 11th edition.

McCallum, B. T. (2011): "Causality, Structure and the Uniqueness of Rational Expectations Equilibria," *The Manchester School*, 79, 551–566, https://doi.org/10.1111/j.1467-9957.2010.02215.x.

McKay, A., E. Nakamura, and J. Steinsson (2016): "The Power of Forward Guidance Revisited," *American Economic Review*, 106, 3133–58, 10.1257/aer.20150063.

——— (2017): "The Discounted Euler Equation: A Note," *Economica*, 84, 820–831, https://doi.org/10.1111/ecca.12226.

Peersman, G. (2005): "What caused the early millennium slowdown? Evidence based on vector autoregressions," *Journal of Applied Econometrics*, 20, 185–207, https://doi.org/10.1002/jae.832.

Ramey, V. (2016): "Chapter 2 - Macroeconomic Shocks and Their Propagation," Volume 2 of Handbook of Macroeconomics: Elsevier, 71–162, https://doi.org/10.1016/bs.hesmac.2016.03.003.

Romer, C. D., and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94, 1055–1084, 10.1257/0002828042002651.

Rubio-Ramírez, J. F., D. F. Waggoner, and T. Zha (2010): "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, 77, 665–696, `https://ideas.repec.org/a/oup/restud/v77y2010i2p665-696.html`.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606, 10.1257/aer.97.3.586.

Stock, J. H., and M. W. Watson (2001): "Vector Autoregressions," *Journal of Economic Perspectives*, 15, 101–115, 10.1257/jep.15.4.101.

TAYLOR, J. B. (1994): "The Inflation/Output Variability Trade-off Revisited," in *Goals, Guidelines, and Constraints Facing Monetary Policymakers*: Federal Reserve Bank of Boston, 21–38.

——— (1999): "The robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the European central bank," *Journal of Monetary Economics*, 43, 655–679, https://doi.org/10.1016/S0304-3932(99)00008-2.

UHLIG, H. (2005): "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, 52, 381–419, https://doi.org/10.1016/j.jmoneco.2004.05.007.

——— (2017): "Shocks, Sign Restrictions, and Identification," in *Advances in Economics and Econometrics*: Cambridge University Press, 95–127.

WOODFORD, M. (2003): *Interest and prices*, Princeton, NJ [u.a.]: Princeton Univ. Press, `http://gso.gbv.de/DB=2.1/CMD?ACT=SRCHA&SRT=YOP&IKT=1016&TRM=ppn+368740471&sourceid=fbw_bibsonomy`.