Thinking about the Economy, Deep or Shallow?*

Pierfrancesco Mei† Lingxuan Wu‡

Job Market Paper

October 13, 2024

Click here for latest version

Abstract

We propose a theory of *shallow thinking* to capture people's limited understanding of the long causal chains involved in the propagation of macroeconomic shocks. Our theory predicts that deeper causal relations have less influence on beliefs. Our estimation suggests that, on average, people understand only about three steps of propagation. In a New Keynesian model with shallow thinking, (i) long-term nominal interest rates overreact to monetary policy shocks but underreact to cost-push shocks; (ii) inflation expectations negatively predict bond excess returns, controlling for yields; (iii) cost-push shocks are more inflationary than under rational expectations; and (iv) more persistent cost-push shocks lead to higher inflation, contrary to the prediction of rational expectations. In a real business cycle model, in response to productivity shocks, (i) output displays a hump-shaped, more persistent response; (ii) investment, labor hours, and the stock price show amplified reactions; and (iii) investment negatively predicts stock excess returns.

*We thank Peter Andre, Adrien Bilal, John Campbell, Juanma Castro-Vincenzi, Gabriel Chodorow-Reich, Ben Enke, Xavier Gabaix, Thomas Graeber, Sam Hanson, Oleg Itskhoki, Spencer Kwon, David Laibson, Chen Lian, Matthew Rabin, Ken Rogoff, Chris Roth, Karthik Sastry, Josh Schwartzstein, Dmitriy Sergeyev, Andrei Shleifer, Stefanie Stantcheva, Jeremy Stein, Ludwig Straub, Adi Sunderam, Alireza Tahbaz-Salehi, Chris Wolf, and seminar participants at Harvard for helpful comments. We thank Sam Cohen for research assistance and Roberto Colarieti for help with the survey design. The study received IRB approval from Harvard (IRB22-1403, IRB24-0959). We acknowledge financial support from the Michael S. Chae Macroeconomic Policy Fund and the Molly and Domenic Ferrante Fund, both awarded through Harvard. Wu is grateful for support from the Alfred P. Sloan Foundation Pre-doctoral Fellowship in Behavioral Macroeconomics, awarded through the NBER.

†Harvard University, pierfrancescomei@g.harvard.edu, https://www.pierfrancescomei.com.

‡Harvard University, lingxuanwu@g.harvard.edu, https://www.lingxuanwu.me, *Corresponding Author*.

1 Introduction

An economy's response to shocks involves causal changes in various prices and quantities, such as inflation, interest rates, and firms' dividends. In general equilibrium models of the economy, prices and quantities emerge from complex causal relations of agents' responses and market outcomes. Expectations of these variables are crucial in macroeconomics and finance because they influence household consumption, firms' pricing and capital investment, asset pricing, and other critical aspects of the economy. How do actual economic agents understand these causal relations when forming their expectations? And how do these expectations, in turn, affect an economy's response to shocks?

The prevailing rational expectations hypothesis amounts to assuming that people fully grasp all causalities in the economy. However, growing evidence suggests that people do not understand the responses of macroeconomic policies to shocks (Bauer, Pflueger and Sunderam, 2024) or the effects of shocks and policy changes on the economy (Andre et al., 2022; D'Acunto, Hoang and Weber, 2022; Coibion et al., 2023*b*). Research from behavioral economics and cognitive psychology further establishes that human reasoning about complex causal systems is limited compared to the benchmark of rationality.[1]

In light of these insights, we propose a theory of *shallow thinking* to model people's limited understanding of the economy as a general equilibrium.[2] We conceptualize a general equilibrium as a set of causal relations in a "directed graph." These causal relations capture how one variable depends on other variables in the economy, resulting from agents' responses and the determination of prices in competitive markets. The rational expectations hypothesis assumes that agents fully comprehend all causal relations in this graph. In contrast, motivated by the aforementioned evidence, we assume that people only understand short chains starting from a shock. As a result, causal relations that are deeper—that is, further removed from shocks—exert less influence on people's beliefs.

We apply our theory to the workhorse New Keynesian and Real Business Cycle (RBC)

[1] People underappreciate how new policies lead to new equilibriums in economic settings (Dal Bó, Dal Bó and Eyster, 2018). They struggle to understand complex causal relations for predictive tasks (Kendall and Oprea, 2024), and make predictions that are insufficiently sensitive to the strengths of causal relations (Rottman and Hastie, 2014). Further, they pay special attention to earlier nodes in causal chains (Ahn et al., 2000), and their knowledge of causally complex systems is sparse and shallow (Rozenblit and Keil, 2002).

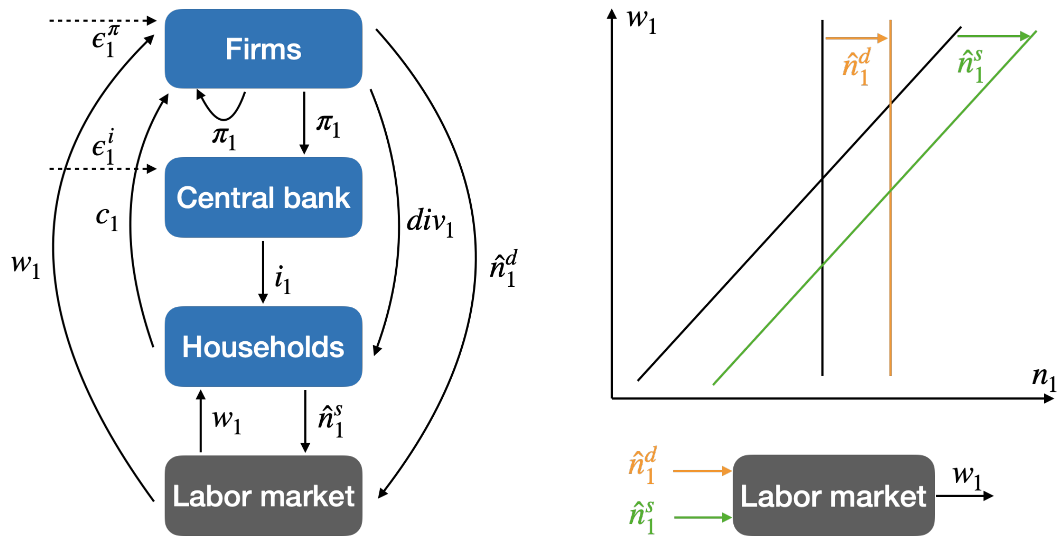
[2] Recent works have relaxed the assumption of full-information rational expectations by removing common knowledge (Angeletos and Lian, 2018) or by modeling agents' limited strategic sophistication (García-Schmidt and Woodford, 2019; Farhi and Werning, 2019) or myopia (Gabaix, 2020), among other notable contributions. However, these studies still assume that agents understand general equilibrium. We will discuss our connection to these works in more detail later.

models and uncover a range of consequences for macroeconomics and finance. We demonstrate that these models feature multiple causal relations that either amplify or offset a shock. By assigning less weight to deeper relations, shallow thinking alters the sign or magnitude of the perceived *net* general equilibrium effect. Consequently, beliefs may over- or underreact to shocks (compared to equilibrium outcomes), depending on the specific causal relations involved. Among other findings, we show that in the New Keynesian model, (i) long-term nominal interest rates overreact to monetary policy shocks but underreact to cost-push shocks, reconciling several bond market puzzles (Cochrane and Piazzesi, 2002; Gürkaynak, Sack and Swanson, 2005; Hanson and Stein, 2015; Bauer, Pflueger and Sunderam, 2024; Joslin, Priebsch and Singleton, 2014; Cieslak, 2018); and (ii) cost-push shocks are more inflationary than under rational expectations. In the RBC model, in response to productivity shocks, (i) output exhibits a hump-shaped, more persistent response, and (ii) investment negatively predicts stock excess returns, consistent with empirical findings of Greenwood and Hanson (2015).

In Section 2, we start by conceptualizing the textbook New Keynesian model à la Woodford (2003*b*) and Galí (2015) as a set of causal relations, and introduce shallow thinking. We study news shocks that are observed in period 0 but only affect the economy in period 1, demonstrating later that the insights hold when generalized to persistent shocks. In period 1, agents observe all variables, respond optimally, and markets clear, making the period-1 economy a static general equilibrium independent of agents' belief formation. However, the period-0 equilibrium depends on agents' beliefs about period-1 outcomes, as households' and firms' decisions are forward-looking. We introduce shallow thinking regarding period-1 outcomes and explore its consequences for the period-0 equilibrium.

Figure 1a illustrates the causal relations of the period-1 economy in a directed graph, subject to a cost-push shock ϵ_1^π and a monetary policy shock ϵ_1^i. Each node indicates a type of agents or a competitive market, and each arrow is an endogenous variable—either a decision of agents or a price. In the case of a competitive market, we interpret the price as set by a fictitious Walrasian auctioneer to equilibrate supply and demand, as demonstrated in Figure 1b. The causal relations among variables are embedded by agents' best responses and the determination of prices in markets. The graph 1a is cyclic, representing the equilibrium as a fixed point. The rational expectations hypothesis assumes that agents take infinite steps in this graph to converge to the fixed point.

In contrast to rational expectations, we hypothesize that each individual only foresees a finite number of steps of shock propagation in the graph, which we refer to as their



(a) Directed graph of causal relations

(b) Competitive labor market

Figure 1: Period-1 New Keynesian economy

individual "depth of thinking" d, and that d varies across the population. Intuitively, when thinking about a cost-push shock ϵ_1^π, a depth-1 agent acknowledges only the most obvious implication: firms will raise their prices (causing inflation π_1), while overlooking changes in all other variables. A depth-2 agent can further appreciate that the central bank will raise the interest rate i_1 in response to higher inflation, but fails to foresee additional implications. A depth-3 agent understands that a higher interest rate i_1 will discourage household consumption and incentivize labor supply. A depth-4 agent recognizes that changes in household behavior will affect the firms and the labor market. This iteration continues infinitely, and only a depth-∞ (i.e., rational) agent correctly assesses the strength of all loops and accurately forecasts the period-1 equilibrium. This iterative process formally captures the idea that deeper implications are harder to grasp.

Our theory makes a novel prediction about belief heterogeneity, which also allows us to measure the distribution of d across the population. An agent with a low d perceives only a subset of variables that are shallow—i.e., close to the shocks in the directed graph. Thus our theory predicts that deeper variables are understood by fewer people.

For tractability, we assume that the depth of thinking d follows a geometric distribution with continuation rate λ. This parameter λ is the only input required to apply our theory to macroeconomic models. A higher λ means that people think more deeply on average, with $\frac{1}{1-\lambda}$ representing the average depth of thinking, and $\lambda = 1$ nesting the rational expectations hypothesis. The average expectations, which drive the aggregate outcomes in a large class of models (including those analyzed in this paper), is tractably parametrized by λ. They are as if generated by a representative who knows all causal relations in the economy but dampen them by λ. Deeper causal effects are dampened exponentially, as they take more steps to reach, and thus exert less influence on beliefs.

In Section 3, we present our survey to test the prediction that deeper variables are understood by fewer people and measure the parameter λ. We use hypothetical vignettes to investigate people's understanding of the propagation of classic macroeconomic shocks, such as oil shocks and monetary policy shocks. For each shock scenario, we ask respondents to provide directional forecasts of changes in a host of important macroeconomic variables, heresuch as inflation and interest rates. We use the directional responses from the empirical literature as the objectively correct answers. We show that variables deeper relative to shocks are indeed forecasted with correct directions by fewer respondents, as predicted by the theory. This finding is robust to including a rich set of fixed effects which absorb sources of belief heterogeneity that do not correlate with depth.

Leveraging the empirical findings, our estimation indicates that λ is around 0.69 [xx subject to change as we enlarge the sample], which implies that the average depth of thinking is only about three—far below infinity (per rational expectations). Furthermore, we find that respondents who better forecast the effects of one shock tend to better understand another shock as well. This suggests that the ability to understand the economy is indeed an individual characteristic, as hypothesized.

In Section 4, we delineate the consequences of shallow thinking for asset prices and aggregate outcomes in the New Keynesian economy, first considering transitory news shocks and then generalizing to persistent shocks. The period-0 long-term interest rate *underreacts* to news about period-1 cost-push shocks, consistent with findings by Bauer, Pflueger and Sunderam (2024), because agents underappreciate the monetary policy reaction in period 1. However, period-0 expectations of period-1 inflation *overreact*. This occurs because, in period 1, the shallower causal relations amplify the inflation response to cost-push shocks, while the deepest relation—monetary policy reaction—offsets it. Since shallow agents better understand the shallow relations, they perceive net amplification, even though the true net effect is offset.

Conversely, the period-0 long-term interest rate *overreacts* to news about period-1 monetary policy shocks, aligning with the findings of a large literature (Cochrane and Piazzesi, 2002; Gürkaynak, Sack and Swanson, 2005; Hanson and Stein, 2015). In period

1, all causal relations offset the effects of monetary policy shocks on the interest rate. As they are dampened to varying degrees, shallow agents perceive less offset than the true effect. As a result, period-0 interest rate expectations, and thus the period-0 long-term interest rate, overreact to monetary policy shocks. However, as agents underappreciate the effect of monetary policy on inflation, inflation expectations *underreact*.

Moreover, variables other than current yields predict bond excess returns when multiple shocks impact the economy, as noted by Joslin, Priebsch and Singleton (2014), Cieslak (2018) and others. In the presence of multiple shocks, interest rate expectations (and thus current yields) primarily reflect monetary policy shocks, while other macroeconomic variables load more on other shocks. This leads to the predictability of bond excess returns by macroeconomic variables, controlling for current yields. Taken together, shallow thinking reconciles several bond market puzzles that seem unrelated or even contradictory.

In terms of aggregate outcomes, news about period-1 cost-push shocks is more inflationary and less contractionary in period 0 than under rational expectations. The underlying mechanism is that, upon receiving news about future cost-push shocks, agents overestimate future inflation and underappreciate the future increase in interest rate and the associated economic contraction. Consequently, in period 0, firms price higher in anticipation of higher future inflation, and households consume more as they expect a lower real interest rate, leading to higher inflation and less contraction in output.

In Section 5, we generalize the analysis to persistent shocks and show that the insights gained from transitory news shocks still hold, while new lessons emerge. A persistent cost-push shock is more inflationary under shallow thinking than under rational expectations, as in the case with transitory news shocks. Additionally, a *more* persistent cost-push shock leads to *higher* inflation under shallow thinking, rather than lower inflation as predicted by rational expectations. In generalizing shallow thinking to accommodate persistent shocks, we focus on causal relations across variables and abstract away from the intertemporal dimension.[3] For example, regarding the dependence of consumption $\{c_t\}_{t\geq 0}$ on interest rates $\{i_t\}_{t\geq 0}$, we assume that if an agent understands how c_t depends on contemporaneous i_t, they also understand how c_t depends on future i_s. With persistent shocks, the causal relations agents underappreciate become the Jacobians of $\{c_t\}_{t\geq 0}$ with respect to $\{i_t\}_{t\geq 0}$, which generalizes the dependence of c_1 on i_1 as a partial derivative in the case of transitory news shocks.

[3] This assumption is for simplicity and generates cross-variable dampening, which complements horizon-dependent dampening in Angeletos and Lian (2018), Farhi and Werning (2019) and Gabaix (2020).

In Section 6, we consider an RBC economy, and demonstrate that shallow thinking amplifies business cycle fluctuations in response to productivity shocks, in a way consistent with return predictability observed in investment boom-bust cycles. In response to a persistent productivity shock, shallow agents believe that firms will invest more and pay higher dividends in the future, but underappreciate that firms' behavior will push up wages in the economy. Due to this underestimation of future wages, firms invest more and hire more labor compared to the rational expectations equilibrium. This overaccumulation of capital leads to a hump-shaped, persistent boom in output and an amplified response in labor hours. Furthermore, since agents overestimate future dividends and underestimate future interest rates, the stock price becomes more volatile. However, as overly optimistic beliefs prove incorrect, realized stock returns are lower. Thus, higher current investment and earnings in this economy predict lower future stock excess returns, consistent with the findings of Greenwood and Hanson (2015).

1.1 Literature Review

At a high level, our theory enriches prior works by suggesting that, among multiple general equilibrium effects, deeper ones are more dampened. Angeletos and Lian (2023*a*) review recent research that moves beyond full-information rational expectations and highlight a key commonality: in a stylized model with one partial equilibrium effect and one general equilibrium (GE) effect, several prominent theories are equivalent in dampening the GE effect. Building on this insight, we show that workhorse macroeconomic models feature multiple GE effects, sometimes with opposing signs. By assigning less weight to deeper ones, shallow thinking alters the sign or magnitude of the perceived *net* GE effect.

Specifically, our theory concerns rationality in the absence of information frictions, and complements a large theoretical literature that studies information frictions (Lucas, 1972; Gabaix and Laibson, 2001, Mankiw and Reis, 2002; Woodford, 2003*a*; Nimark, 2008; Angeletos and Lian, 2018), rational inattention (Sims, 2003; Maćkowiak and Wiederholt, 2009; Molavi, 2019; Miao, Wu and Young, 2022) and learning (Evans and Honkapohja, 2001; Eusepi and Preston, 2018). Our survey studies understanding of macroeconomic shocks under full information and shows that people fail to even get the directions of impulse responses correct, which provides unique support to our theory.

Our theory broadens an important line of research that considers agents' limited strategic sophistication in macroeconomics (García-Schmidt and Woodford, 2019; Farhi and

Werning, 2019; Iovino and Sergeyev, 2023; Bianchi-Vimercati, Eichenbaum and Guerreiro, 2024) and finance (Greenwood and Hanson, 2015; Bastianello and Fontanier, 2024). These works introduce models of level-k thinking (Nagel, 1995; Stahl and Wilson, 1994, 1995; Camerer, Ho and Chong, 2004) and competition neglect (Camerer and Lovallo, 1999) from the experimental and game-theoretical literature to general equilibrium. Level-k thinking addresses agents' limited strategic sophistication when lacking experience in analogous games. The aforementioned papers aptly apply it to study unconventional macroeconomic policies (such as forward guidance and quantitative easing) in New Keynesian models. Shallow thinking accounts for agents' underappreciation of non-strategic causal relations (such as the monetary policy rule) and applies to models that are not typically considered strategic (such as the RBC model), as well as to conventional shocks, all driven by a lack of knowledge. Moreover, in our survey, we measure strategic sophistication using the classic "guess 2/3 of the average" game, and find *no* correlation with understanding of macroeconomic shocks. A reasonable interpretation is that shallow thinking reflects limited knowledge about the macroeconomy—a different aspect of bounded rationality from limited strategic sophistication.

Our theory focuses on the dampening of causalities *across variables*, which complements *horizon-dependent* dampening due to bounded rationality (Gabaix, 2020; Farhi and Werning, 2019) or information frictions (Angeletos and Lian, 2018; Angeletos and Huo, 2021). With persistent shocks, we assume that if an agent understands how one variable contemporaneously depends on another variable, they also understand how it depends on future values of the other variable. One could generalize our theory to accommodate horizon-dependent dampening by introducing failure of causal reasoning across time.

Our theory generates belief over- and underreaction in a manner endogenous to the causal relations involved. It implies belief overreaction of a variable to shocks that hit itself, if the very indirect general equilibrium effect is offset, but underreaction otherwise. Hence it adds a layer of richness to theories of overreaction (Barberis, Shleifer and Vishny, 1998; Bordalo et al., 2020; Afrouzi et al., 2023; da Silveira, Sung and Woodford, 2024), in a way that reconciles several bond market puzzles regarding over- and underreaction of long-term interest rates (Hanson and Stein, 2015; Bauer, Pflueger and Sunderam, 2024; Joslin, Priebsch and Singleton, 2014; Cieslak, 2018), as previously discussed.

Last, this paper provides a theory of heterogeneous mental models that connect to a growing literature that uses surveys to measure people's mental models in specific scenarios (Stantcheva 2021, 2023; Andre et al. 2022, 2024; Andre, Schirmer and Wohlfart

2024). In particular, Andre et al. (2022) show that people's level forecasts of unemployment and inflation in response to macroeconomic shocks are highly heterogeneous. Our theory predicts when people can or cannot get the directions correct, bringing some order to their findings of belief heterogeneity. We design a survey to confirm our prediction and calibrate a structural parameter of beliefs that could be used in macroeconomic models. Prior to this paper, Wu (2023) calibrates people's imperfect mental models using existing survey forecasts. This paper develops a theory-informed survey to offer additional evidence in its support, and derives its consequences for macroeconomics and finance.

2 Shallow Thinking in a New Keynesian Economy

We set up the textbook New Keynesian model in Section 2.1 with transitory news shocks, which are observed in period 0 but only affect the economy in period 1. We conceptualize the period-1 equilibrium as a set of causal relations in a directed graph in Section 2.2. Accordingly, in Section 2.3, we introduce the concept of shallow thinking and present its main testable prediction: deeper variables—those further removed from shocks—are understood by fewer people. Later in the paper, we examine the consequences of shallow thinking for the period-0 equilibrium and generalize to persistent shocks.

2.1 The New Keynesian Economy

We consider the New Keynesian model à la Woodford (2003*b*) and Galí (2015). The economy consists of three types of agents (firms, households, and a central bank) and a competitive labor market. Throughout this paper, we take a log-linear approximation around the steady state for simplicity and use lower-case letters for log-linear deviations.

We study news shocks that are observed in period 0 but only affect the economy in period 1. Since this New Keynesian model is purely forward-looking, the economy returns to its steady state from period 2 onwards. Appendix A.1 develops the infinite-horizon model in full detail. Here, we focus on the period-1 general equilibrium, and conceptualize it as a set of causal relations purposefully as follows: (i) we maintain the structural form of agents' best responses, which express their optimal decisions as functions of decision-relevant variables, and (ii) we interpret price determination in a competitive market as a rule that pins down the price to equilibrate supply and demand, with a fictitious Walrasian auctioneer. These forms are central to our theory.

Firms. There is a continuum of firms indexed by $j \in [0,1]$ that produce using labor to satisfy demand and set prices subject to Calvo rigidity. In period 1, firms choose labor demand, pay dividends, and reset prices if possible, taking as given the aggregate inflation rate π_1, the real wage w_1, and the aggregate demand c_1. Each firm produces a differentiated good, which collectively forms a constant-elasticity bundle that households consume, and charges a markup μ in the steady state.

All firms produce to satisfy demand using the same linear technology in labor, giving rise to the aggregate dividend and labor demand

$$
\begin{aligned}
div_1 &= c_1 - \frac{1}{\mu - 1} w_1 \\
n_1^d &= c_1
\end{aligned}
\tag{1}
$$

To anticipate our analysis of the labor market, we interpret labor demand n_1^d as a demand curve $n_1^d = \hat{n}_1^d + E_{n^d w} w_1$, which shifts by

$$
\hat{n}_1^d = c_1 \tag{2}
$$

and has an elasticity $E_{n^d w}$, which is zero in this model, as firms only use labor in production.

A $(1-\theta)$ share of firms can reset their prices in period 1 to maximize dividends, and each chooses

$$
p_{j1}^* = p_0 + (1 - \beta\theta) \left[w_1 + \sum_{k=0}^{\infty} (\beta\theta)^k \pi_1 \right]
$$

where β is the household time discount rate, the inverse of which equals the steady-state interest rate. Aggregate inflation results from the pricing behavior of the $(1-\theta)$ share of resetting firms as $\pi_1 = (1-\theta)\left(p_{j1}^* - p_0\right)$. Following the tradition, we consider a cost-push shock ϵ_1^π, and thus inflation is

$$
\pi_1 = \theta\kappa w_1 + (1-\theta)\,\pi_1 + \epsilon_1^\pi \tag{3}
$$

with $\kappa \equiv \frac{(1-\theta)\left(1-\beta\theta\right)}{\theta}$ capturing the slope of the Phillips curve. Importantly, we do not move π_1 on the right-hand side to the left. We intentionally preserve the dependence of π_1 on itself, which encapsulates the within-period complementarity in individual price-setting, as each firm takes the aggregate inflation as given.

Households. There is a continuum of infinitely lived households who maximize their lifetime utility, discounted by β, which is separable in consumption and labor supply. In period 1, households choose consumption and labor supply, taking as given the nominal interest rate i_1, the real wage w_1, and dividend div_1. Their optimal consumption and labor supply decisions are given by

$$c_1 = -\sigma^{-1}\beta i_1 + \frac{(1-\beta)(\mu-1)\nu}{\sigma+\mu\nu}div_1 + \frac{(1-\beta)(1+\nu)}{\sigma+\mu\nu}w_1 \tag{4}$$

$$n_1^s = \nu^{-1}\beta i_1 - \frac{(1-\beta)(\mu-1)\sigma}{\sigma+\mu\nu}div_1 + \nu^{-1}\left[1-\sigma\frac{(1-\beta)(1+\nu)}{\sigma+\mu\nu}\right]w_1$$

where σ^{-1} is the elasticity of intertemporal substitution, and ν^{-1} is the the Frisch elasticity of labor supply.

Similar to labor demand, we interpret labor supply n_1^s as a supply curve $n_1^s = \hat{n}_1^s + E_{n^s w}w_1$, which shifts by

$$\hat{n}_1^s = \nu^{-1}\beta i_1 - \frac{(1-\beta)(\mu-1)\sigma}{\sigma+\mu\nu}div_1 \tag{5}$$

and has an elasticity $E_{n^s w} = \nu^{-1}\left[1-\sigma\frac{(1-\beta)(1+\nu)}{\sigma+\mu\nu}\right]$.

Central bank. The central bank follows a Taylor rule with a monetary policy shock ϵ_1^i,

$$i_1 = \phi\pi_1 + \epsilon_1^i \tag{6}$$

Competitive labor market. Finally, to close the model, the wage arises from equilibrating labor supply and demand $n_1^s = n_1^d$. We introduce a fictitious labor market auctioneer who determines the wage as the intersection of supply and demand curves,

$$\hat{n}_1^s + E_{n^s w}w_1 = \hat{n}_1^d + E_{n^d w}w_1$$

$$w_1 = (E_{n^s w} - E_{n^d w})^{-1}\left(\hat{n}_1^d - \hat{n}_1^s\right) \tag{7}$$

which prescribes that the wage is higher if labor demand is higher or labor supply is lower, as illustrated in Figure 1b.

2.2 General Equilibrium as Causal Relations

We interpret the period-1 equilibrium as a set of causal relations in a directed graph.

The period-1 equilibrium in response to the cost-push shock ϵ_1^π and the monetary policy shock ϵ_1^i is fully characterized by (1-7). We collect all endogenous variables in a vector $V_1 \equiv \left(i_1, \pi_1, div_1, \hat{n}_1^d, c_1, \hat{n}_1^s, w_1\right)'$ and the two shocks correspondingly in $S_1 \equiv \left(\epsilon_1^i, \epsilon_1^\pi, 0, 0, 0, 0, 0\right)'$. While we focus on these two shocks in this paper, it is straightforward to incorporate additional shocks.

We rewrite (1-7) as a *fixed point* system in matrix form

$$\underbrace{V_1}_{\text{variables}} = \underbrace{M}_{\text{causal relations}} V_1 + \underbrace{S_1}_{\text{shocks}} \tag{8}$$

where M is a matrix capturing all the partial derivatives among the endogenous variables. Each partial derivative represents a causal relation. Each equation in this system describes how the outcome on the left responds to a set of causes on the right.

We visualize this causal system as a directed graph in Figure 1a, formally supporting the intuition outlined in the introduction. Each node represents a type of agent (including the fictitious labor market auctioneer), and each arrow indicates an endogenous variable—either an agent's decision or the real wage determined in the labor market.

We solve this system (8) to express the equilibrium as a *sum* of all effects

$$V_1 = (I - M)^{-1} S_1 = \sum_{n=1}^{\infty} M^{n-1} S_1 \tag{9}$$

where each M^{n-1} term is an n-step effect of a shock on a variable via $n-1$ variables. The 1-step effect S_1 is the direct (or partial equilibrium) effect of shocks, while all subsequent terms represent indirect (or general equilibrium) effects.

As these period-1 shocks are observed in period 0, if agents are rational, they will correctly forecast the period-1 equilibrium, i.e., $\mathbb{E}_0^{rational}[V_1] = V_1$. In this sense, the rational expectations hypothesis assumes that agents can take infinite steps in this graph to converge to the fixed point. Next, we introduce our theory of shallow thinking.

2.3 Shallow Thinking

Motivated by evidence in economics and psychology, we assume that agents only foresee finite steps of shock propagation in the directed graph. We outline the key assumptions and derive the testable prediction that deeper variables are perceived by fewer people.

Assumption 1. Individuals vary in their finite *depth of thinking* $d \in \mathbb{N}^+$,

$$\mathbb{E}_0^d [V_1] \equiv \sum_{n=1}^{d} M^{n-1} S_1 \tag{10}$$

A depth-d agent only understands the effects that take no more than d steps. Their beliefs also satisfy an iterative formula as

$$\mathbb{E}_0^d [V_1] = M\mathbb{E}_0^{d-1} [V_1] + S_1 \tag{11}$$

which captures the intuition that a depth-d agent can think one more step compared to a depth-($d-1$) agent.

For example, in response to a cost-push shock ϵ_1^π (Figure 1a), a depth-1 agent acknowledges only the most obvious implication: firms will raise their prices (causing inflation π_1). A depth-2 agent can further appreciate that the central bank will raise the interest rate i_1 in response to higher inflation. A depth-3 agent understands that a higher interest rate i_1 will discourage household consumption and incentivize labor supply. A depth-4 agent recognizes that changes in household behavior will affect the firms and the labor market. This iteration continues infinitely, and only a depth-∞ (i.e., rational) agent correctly assesses the strength of all loops and accurately forecasts the period-1 equilibrium.

This iterative process leads to a prediction about belief heterogeneity, which allows us to measure the distribution of d across the population shortly. We observe that agents with a low d only perceive variables that are shallow—that is, closer to shocks in the directed graph. And the set of variables that they can perceive varies with shocks. By examining people's expectations of various macroeconomic variables in response to different shocks, we can measure the distribution of d, which we formally establish shortly.

For tractability, we make an additional parametric assumption.

Assumption 2. Individual depth of thinking d follows a geometric distribution over $\mathbb{N}^+$

with continuation rate $\lambda \in [0, 1]$, i.e.,

$$\mathbb{P}(d \geq n) = \lambda^{n-1},\ \forall n \in \mathbb{N}^{+} \tag{12}$$

A higher λ means that individuals are deeper on average, with $\frac{1}{1-\lambda}$ representing the average depth of thinking, and $\lambda = 1$ nesting the rational expectations hypothesis.. With this parametric assumption, we could aggregate heterogenous beliefs into average beliefs, which will drive the period-0 equilibrium.

Proposition 1. (Average beliefs) *The average beliefs* $\overline{\mathbb{E}}_0[V_1] \equiv \sum_{n=1}^{\infty} \mathbb{P}(d = n) \cdot \mathbb{E}_0^n[V_1]$ *are sums of all effects*

$$\overline{\mathbb{E}}_0[V_1] = \sum_{n=1}^{\infty} \lambda^{n-1} M^{n-1} S_1 \tag{13}$$

and equivalently, as a fixed point

$$\overline{\mathbb{E}}_0[V_1] = \underbrace{\lambda M}_{\text{average perceived causal relations}} \overline{\mathbb{E}}_0[V_1] + S_1 \tag{14}$$

Equation (13) is comparable to (9) that expresses the equilibrium as a sum of all effects, but with deeper effects dampened more, since fewer people appreciate them. Equation (14) further suggests that, the average beliefs as a fixed point are formed as if by a representative agent who knows all causal relations in M but underappreciates them by a factor λ, parallel to (8) with the equilibrium as a fixed point. The proof is simple, by summing all n-step effects of shock propagation with decaying weights.

Proof of Proposition 1. Assumptions 1 and 2 imply that the average beliefs are

$$\begin{aligned}
\overline{\mathbb{E}}_0[V_1] &\equiv \sum_{d=1}^{\infty} \mathbb{P}(d = n)\, \mathbb{E}_0^d[V_1] \\
&= \sum_{d=1}^{\infty} \mathbb{P}(d = n) \sum_{n=1}^{d} M^{n-1} S_1 = \sum_{n=1}^{\infty} \mathbb{P}(d \geq n)\, M^{n-1} S_1 = \sum_{n=1}^{\infty} \lambda^{n-1} M^{n-1} S_1
\end{aligned}$$

which can be recast as (14).□

Moreover, equation (14) coincides exactly with the formulation of imperfect mental model in Wu (2023). That paper extracts an empirical moment based on this formula using existing forecasts data and rejects the null of $\lambda = 1$ (rational expectations). The nature of that exercise is *quantitative*, as it compares forecasts to the true conditional expectations. In this paper, we provide *qualitative* evidence to support our theory based on heterogeneity in beliefs being directionally correct, elicited in a customized survey introduced next.

We design a survey to elicit people's understanding of macroeconomic shocks. We recruit a representative sample of respondents indexed by n. We ask each respondent to provide *directional* forecasts of a set of macroeconomic variables v, in response to different hypothetical shocks s, over a 12-month horizon. We get the correct direction of change in each variable v in response to each shock s from the empirical literature, and determine accordingly whether each respondent n's directional forecast is correct.

Definition 1. We define *correct directional belief* 1_{nvs} as one if respondent n correctly forecasts the directional response of variable v to shock s and zero otherwise.

As discussed earlier, an agent with a low depth of thinking d only perceives causal relations and variables close to shocks. When aggregated across the population, a variable further removed from a shock is understood by fewer people. This is a prediction at the population level, without the need to determine the depth of thinking d for each individual, which facilitates our empirical test.

In order to formally define the depth of a variable relative to a shock for our test, we introduce some additional notations. Notice that beliefs $\mathbb{E}^d[V_1]$ are linear in the shocks S_1, as determined in (10). With slight abuse of notation, we let $v \in V_1$ be an endogenous variable in our model and $s \in S_1$ be a shock. Thus, $\frac{\partial v}{\partial s}$ is the true sensitivity of variable v to shock s, whereas $\frac{\partial \mathbb{E}^d[v]}{\partial s}$ is the perceived sensitivity by a depth-d individual.

Definition 2. We define the *variable depth* D_{vs} as the minimum d such that $\frac{\partial \mathbb{E}^d[v]}{\partial s}$ has the same sign as $\frac{\partial v}{\partial s}$, for each variable $v \in V_1$ and shock $s \in S_1$.

That is, variable depth D_{vs} corresponds to the depth of the shallowest individual who can correctly perceive the directional response of v to s. In our example with transitory cost-push and monetary policy shocks, D_{vs} equals the depth of the shallowest agent who perceives any change of v in response to s. That is, $\frac{\partial \mathbb{E}^d[v]}{\partial s}$ is zero for all $d < D_{vs}$. Nonetheless, Definition 2 is more general when applied to persistent shocks and other models.

Assumption 3. Model parameters M are such that $\frac{\partial \mathbb{E}^d[v]}{\partial s}$ has the same sign as $\frac{\partial v}{\partial s}$ for all $d \geq D_{vs}$.

Assumption 3 holds true when deeper effects either amplify or offset the impact of the shock, once the correct direction is established, but *do no overturn it.* It facilitates a reduced-form estimation of λ as we establish next. This assumption is generically true in the New Keynesian model. For instance, in response to the cost-push shock ϵ_1^π, the central bank will raise the interest rate i_1 to offset the shock, but does not lead to deflation. That is, $\frac{\partial \mathbb{E}^d[\pi_1]}{\partial \epsilon_1^\pi}$ is positive for all $d \geq 1$. And since $\mathbb{E}^d[i_1] = \phi \mathbb{E}^{d-1}[\pi_1]$ from (11), $\frac{\partial \mathbb{E}^d[i_1]}{\partial \epsilon_1^\pi}$ is positive for all $d \geq 2$. Further, even if it is not true for all variable-shock combinations, as long as there exists a subset of such combinations with varying D_{vs}, our estimation can go through by focusing on this subset.

Proposition 2. (Heterogeneity in correct directional beliefs) *The expectation of correct directional belief 1_{nvs}, conditional on variable depth D_{vs}, in the population is*

$$\mathbb{E}^{pop}[1_{nvs}|D_{vs} = D] = \lambda^{D-1}, \ \forall D \in \mathbb{N}^+ \tag{15}$$

where $\mathbb{E}^{pop}$ denotes the expectation in the population of survey respondents. Consequently,

1. *an ordinary least squares estimation of $1_{nvs} = \gamma D_{vs} + \alpha + \epsilon_{nvs}$ yields a negative slope β;*
2. *a nonlinear least squares estimation of $1_{nvs} = b_1 \cdot b_2^{D_{vs}-1} + b_0 + \epsilon_{nvs}$ identifies λ with b_2.*

We consider both the nonlinear and the linear specifications. The null of $\gamma = 0$ and $b_2 = 1$ includes rational expectations and any other theories of beliefs that do not correlate with variable depth D_{vs}. Our estimation result in Section 3 will show that γ is negative and b_2 is lower than 1, both with high levels of statistical significance.

The nonlinear specification lets us estimate λ from a regression. (15) suggests that, under our three assumptions, the conditional expectation of 1_{nvs}, which is the conditional probability of making correct directional forecasts, is exponentially decaying. Thus a nonlinear least-squares estimation of an exponential function can exactly recover λ. Our Assumption 3 crucially facilitates this estimation. As discussed earlier, if Assumption 3 does not hold for all possible combinations of variables and shocks, as long as one can find a subset of such combinations with varying D_{vs}, one can still estimate λ with the nonlinear regression on this subset. If even that is not possible, one can estimate λ by minimizing distance between the distribution of measured 1_{nvs} and the corresponding theory-implied distribution, as λ parametrizes the latter distribution.

A linear specification is valuable for two reasons: (i) it allows us to empirically control for fixed effects to purge confounding sources of belief heterogeneity, and we will show that our coefficient of interest γ is indeed robust to such controls; and (ii) it does not hinge on the parametric Assumption 2. A negative γ by itself indicates that some agents only understand variables close to shocks. When Assumption 2 does hold, the estimated slope γ will be a weighted average of the local slopes of the nonlinear function λ^{D-1}.

In summary, the idea of shallow thinking is that people understand only limited steps of shock propagation, captured by Assumption 1 as the backbone of our theory. Assumption 2 serves as a convenient aggregator to generate average beliefs. The nature of Assumption 2 is parametric rather than conceptual, akin to how Calvo pricing is a useful parametrization of nominal rigidity but not essential. With these two assumptions, one could generate heterogeneous and average beliefs with one parameter λ in a macroeconomic model. If the model satisfies Assumption 3 in addition, the variable depth D_{vs} can be used as a predictor of belief heterogeneity to test shallow thinking and estimate λ.

3 Measuring Shallow Thinking

We develop a survey to test the theoretical prediction of shallow thinking that deeper variables are understood by fewer people (Proposition 2). We offer further evidence in Section 3.2 to suggest that the limited depth of thinking vis-à-vis economic causal relations is an individual characteristic and reflects limited knowledge about the macroeconomy.

Survey design. As alluded to before, we elicit people's understanding of classic macroeconomic shocks, by asking them to forecast directional responses of important macroeconomic variables. We study six shocks in three groups: oil price shock (oil) and monetary policy shock (MP) as group 1, government spending shock (G) and personal income tax shock (PIT) as group 2, and corporate income tax shock (CIT) and transfer payment shock (TP) as group 3. Table 1 lists the baseline specification with eleven macroeconomic variables, their correct directional responses and their variable depths D_{vs}. The baseline variable depth D_{vs} is generated from enriching our New Keynesian model with decreasing-returns production and Taylor rule dependent on both inflation and unemployment. We consider various robustness versions regarding the selection of variable-shock combinations (v, s) and variable depth D_{vs} in Tables C3 and C4. We randomly allocate half of the respondents into group 1, which has the two shocks we analyze in the model, and allocate a quarter into groups 2 and 3 respectively for additional evidence. Appendix **??** discusses our survey design in greater detail.

Table 1: Baseline version of variable depth and correct directions

	Group 1 (50%)		Group 2 (25%)		Group 3 (25%)	
	Oil ↑	MP ↑	G ↑	PIT ↑	CIT ↑	TP ↑
Output	3↓	2↓	1↑	2↓	3↓	2↑
Interest rate	2↑	1↑				
Price	1↑	3↓	2↑	3↓		
Unemployment	3↑	2↑	2↓	2↑	3↑	3↓
Labor hours	3↓	2↓	2↑	2↓	3↓	3↑
Durable consumption	3↓	2↓	2↓	2↓	3↓	2↑
Non-durable consumption	3↓	2↓	2↓	2↓	3↓	2↑
Post-tax profits					1↓	1↑
Personal income tax			1↑	1↑		1↑
Corporate income tax			1↑		1↑	1↑
Government borrowing			1↑	1↓	1↓	1↑

Notes: Six shocks are oil price shock (oil), monetary policy shock (MP), government spending shock (G), personal income tax shock (PIT), corporate income tax shock (CIT), and transfer payment shock (TP). The latter four all concern the federal government. Each cell indicates the variable depth D_{vs} and the directional response (up or down) in the baseline specification. The directional responses are from the empirical literature reviewed in Table C2, and robustness versions of selection of variable-shock combinations and variable depth are in Tables C3 and C4.

3.1 Fact 1: Variable Depth Explains Correct Directional Belief

We examine the predictability of correct directional belief 1_{nvs} by variable depth D_{vs} as prescribed by Proposition 2.

Figure 2 shows the expectation of correct directional belief 1_{nvs}, conditional on variable depth D_{vs}. The blue dot indicates the conditional expectation, together with 99.9% confidence interval. The red diamond represents the conditional expectation, after controlling for individual-by-variable and individual-by-shock fixed effects δ_{nv}, δ_{ns}, the purpose of which we discuss soon, together with 99.9% confidence interval. Obviously, the conditional expectation declines in variable depth in both cases.

With this visualization, we make two remarks. First, the conditional expectation of 1_{nvs} being low for depth-2 variables and even lower for depth-3 variables suggests that people on average are bad at understanding macroeconomic shocks. They fail to perceive even the true directional responses of important variables. This paper focuses on belief heterogeneity arising from people's shallow understanding of causal relations,

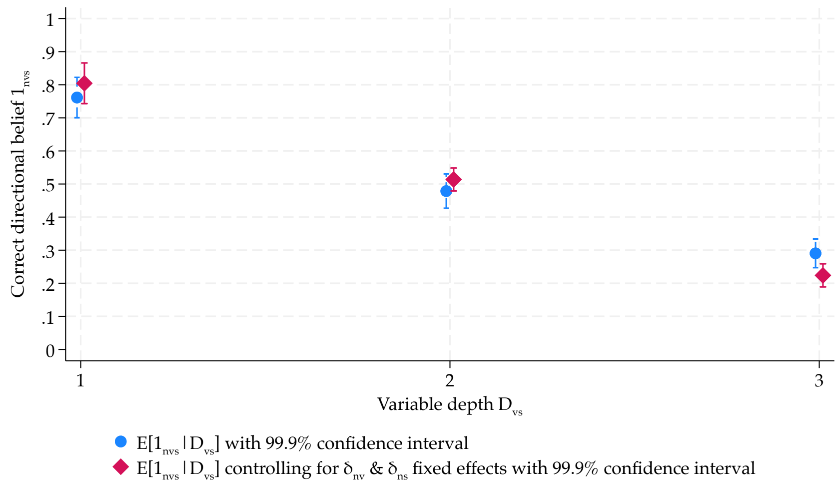


Figure 2: Expectation of correct directional belief 1_{nvs} conditional on variable depth D_{vs}

but assumes for simplicity that they use the true causal structure and model parameters. Figure 2 suggests that our focus is an important aspect, if not the main one, in the following sense. If people can iterate a model infinitely many times, but they believe in alternative causal models, then these models must be quite wrong, and wrong in a way that correlates with variable depth from the New Keynesian model to produce Figure 2. Another possibility is that people iterate a model infinitely many times, but disagree on the model parameters (such as the slope of the Phillips curve). However, those agents typically do not get the directional responses wrong, which is what Figure 2 finds.

Second, when we run the regressions on variable depth, the key parameter λ will not be solely identified off the comparison of depth-1 variables against other variables. One could intuitively expect such a difference, since depth-1 variables are directly shocked but all other variables are only indirectly affected through general equilibrium effects. Our theory further differentiates among the indirectly affected variables by their depth. This empirical finding substantiates our high-level contribution relative to Angeletos and Lian (2023*a*) that among general equilibrium effects, some are better understood than others.

Table 2 presents various specifications of the linear regression

$$1_{nvs} = \gamma D_{vs} + \alpha + \delta_{nv} + \delta_{ns} + \epsilon_{nvs} \tag{16}$$

and the nonlinear regression

$$1_{nvs} = b_1 \cdot b_2^{D_{vs}-1} + b_0 + \epsilon_{nvs} \tag{17}$$

as prescribed by Proposition 2, where 1_{nvs} is one if individual n's directional forecast of variable v in response to shock s is correct.

The coefficient γ from the linear regression tests for the theory-implied pattern that deeper variables are understood by fewer people.[4] The null of $\gamma = 0$ includes rational expectations and any other theory of beliefs that does not correlate with depth D_{vs}. Further, if respondents are totally clueless about the economy and give random answers in our survey, that will not be reflected in γ. Thus, a negative γ not only implies that people make mistakes, but they do so in a depth-dependent way as predicted by our theory.

Table 2: Regression of correct directional belief 1_{nvs} against variable depth D_{vs}

	OLS					NLS
Correct directional belief 1_{nvs}	(1)	(2)	(3)	(4)	(5)	(6)
Variable depth D	-0.22***	-0.25***	-0.26***	-0.29***		
	(0.01)	(0.01)	(0.01)	(0.01)		
$1_{D=2}$					-0.29***	
					(0.03)	
$1_{D\geq 3}$					-0.58***	
					(0.02)	
$b_1 - 1$						-0.12
						(0.21)
$b_2 - 1$						-0.31***
						(0.09)
Observations	5436	5436	5436	5436	5436	5436
R^2	0.10	0.24	0.30	0.63	0.63	0.11
Individual FE		Yes	Yes	Absorbed	Absorbed	
Variable FE			Yes	Absorbed	Absorbed	
Shock FE			Yes	Absorbed	Absorbed	
Individual-variable FE				Yes	Yes	
Individual-shock FE				Yes	Yes	

Standard errors in parentheses
Standard errors clustered at individual level
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Column (1) uses variable depth D_{vs} as the only predictor and finds a statistically significant coefficient with a R^2 of 10%. Column (2) shows that individual fixed effects

[4] Interestingly, in a very different context, using network data of relationships from Indian villages, Breza, Chandrasekhar and Tahbaz-Salehi (2018) show that the knowledge of whether certain pairs of households are linked declines steeply in the pair's network distance to the respondent. While our finding is of a similar form, the distance in our context is distinct—it is a conceptual one about how easy it is for people to associate a variable to a shock.

matter too, increasing the R^2 to 24%. That means some people are more likely to be correct than others, as our theory postulates.

Column (3) shows that the estimate of γ and its statistical significance are robust to the inclusion of variable and shock fixed effects. Controlling for these additional fixed effects addresses a concern that people may understand some variables or some shocks better in a way that happens to correlate with depth. For the same variable, it is understood more poorly when it is further away from a shock. A related concern is that some shocks (like monetary policy shocks) may take longer to transmit into the economy or some variables may be slower in responding, and thus people predict no changes in a fixed horizon. These are absorbed by shock and variable fixed effects too.[5]

Column (4) further controls for individual-by-variable or individual-by-shock fixed effects. They absorb belief heterogeneity that is unrelated to depth. For example, if a person believes in a post-pandemic quantity-constrained model of the economy, they will predict that prices respond to all shocks but quantities are fixed. Another person can believe in a price-constrained model, but to the extent that they do not correlate with depth, such heterogeneity is absorbed by individual-by-variable fixed effects. Individual-by-shock fixed effects absorb the possibility that one person only understands monetary policy shocks whereas another person only understands oil shocks.

Column (5) demonstrates that, relative to depth-1 variables (that are directly shocked), depth-2 variables are understood by fewer people, and depth-3-and-above variables by even fewer. This is what we observe from Figure 2, lending further support to the predicted depth-dependent pattern.

Last, column (6) shows the nonlinear estimation and strongly rejects the null of $b_2 = 1$. The estimation suggests that $\hat{\lambda} \approx 0.69$. That means, people on average only understand about three steps, even under our parametric assumption that there is a distribution of people who could reason more than three steps. We will use this value when we apply our theory in workhorse macroeconomic model in Sections 4 and 6.

We summarize our finding as follows, and Appendix C.2 provides additional results and robustness checks.

Fact 1. *Deeper variables are correctly forecasted by fewer people, and this slope is robust to the inclusion of a rich set of fixed effects. Variable depth and individual fixed effects together explain a quarter of the total variation of directional beliefs being correct.*

[5] We also note that most variables we study do have statistically significant impulse responses to shocks at the 12-month horizon we set in the survey.

3.2 Fact 2: Depth of Thinking is Individual Characteristic and Domain-Specific

[xx This fact 2 will not be part of my job talk so that I can go to consequences more quickly. It can be moved to the appendix of the draft as well.]

We further show that the ability to understand shock propagation is indeed an individual characteristic and does not correlate with a classic measure of strategic sophistication.

The previous fact that individual fixed effects matter for correct beliefs already suggests that some people get more variables correct than others. To further investigate this, we measure individual n's overall understanding of shock s by a *total depth score (TDS)* as

$$TDS_{ns} \equiv \sum_{v} D_{vs} \cdot 1_{nvs} \tag{18}$$

To receive a higher TDS, a respondent needs to get more variables correct and especially deeper variables correct.[6]

To the extent that depth is an individual characteristic as we postulate, we expect each respondent's TDSs to correlate strongly across shocks. To test this, we rank the TDSs from the lowest to the highest for each shock, and correlate the two TDS rank measures across individuals in Table 3. Column (1) confirms this prediction.

In contrast, column (2) suggests that TDS does not correlate with a classic measure of strategic sophistication (level k), via a "guess 2/3 of the average" game we play with survey respondents. This connects to findings in the macroeconomic literature that the measured level k does not predict differential consumption response to inflation news by Dutch households (Coibion et al., 2023*a*) or first- and higher-order inflation expectations of New Zealand firm managers (Coibion et al., 2021).[7] We remark that shallow thinking likely reflects people's limited knowledge about the macroeconomy, a distinct aspect of bounded rationality from limited strategic sophistication. After all, a chess master who could anticipate opponents well may not know macroeconomics, and vice versa.

[6]This TDS is a more robust measure to noise than the depth of the deepest variable that is understood correctly, as we have several variables for each depth and respondents may coincidentally get some correct.

[7]In the experimental literature, Dal Bó, Dal Bó and Eyster (2018) show that voters prefer policy changes that bring in direct benefits but induce larger indirect costs, but their voting behavior is not correlated with level k. Georganas, Healy and Weber (2015) study stability of level k using two families of games: beauty contest games à la Nagel (1995) and undercutting games similar to Arad and Rubinstein (2012). They find that the participants' levels are consistent within the beauty contest family, but do not correlate within the undercutting game family or across two families.

Table 3: Total depth score of one shock against that of the other shock

Total depth score of 2nd shock (rank)	(1)	(2)	(3)	(4)	(5)	(6)
Total depth score of 1st shock (rank)	0.21***					0.19***
	(0.05)					(0.05)
Level k (rank)		-0.05			-0.04	-0.02
		(0.06)			(0.05)	(0.05)
Financial literacy score (rank)			0.26***		0.25***	0.23***
			(0.05)		(0.05)	(0.05)
Education (rank)			0.01		-0.04	-0.06
			(0.06)		(0.06)	(0.06)
Net asset (rank)				0.15*	0.15*	0.16**
				(0.06)	(0.06)	(0.06)
Income (rank)				0.04	0.03	0.04
				(0.06)	(0.06)	(0.06)
Constant	0.39***	0.54***	0.34***	0.41***	0.31***	0.21***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.05)	(0.06)
Observations	319	319	319	319	319	319
R^2	0.05	0.00	0.08	0.03	0.11	0.14

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The rest of the columns suggest that a measure of financial literacy, following Lusardi and Mitchell (2011), but not general education, correlates with TDS. Households' net asset positions also correlate with TDS. The former finding corroborates the idea that shallow thinking reflects people's economic knowledge, for which general education may be too noisy a proxy. These individual characteristics that correlate with TDS remain significant when analyzed together.

We summarize as follows.

Fact 2. *Depth of thinking in the context of macroeconomic shock propagation shows a strong correlation across different shock scenarios for the same participant, indicating it as an individual characteristic. Further, it does not correlate with a classic measure of strategic sophistication, suggesting that it is specific to the macroeconomic domain.*

4 Consequences of Transitory News Shocks

We discuss belief over- and underreaction to shocks due to shallow thinking, and the consequences for asset prices and aggregate outcomes in the New Keynesian economy, in the case with transitory news shocks that are observed in period 0 but only affect the economy in period 1, as introduced in Section 2.

We make an important remark on the generality of analyzing transitory news shocks: while we compare shallow thinking against rational expectations regarding news about period-1 shocks, the same comparison holds for persistent shocks that materialize in period 0 and last for two periods. That is simply because in the log-linearized economy, a 2-period persistent shock is equivalent to the sum of a period-0 shock and a period-1 shock that is observed in period 0. The economy's response to a period-0 shock is independent of agents' belief formation. Thus the comparison across theories of expectations regarding any 2-period persistent shock is solely driven by its news shock component.

Table 4: Quarterly calibration of the New Keynesian model

Parameter	Description	Value	Estimate/Target
Beliefs			
λ	Continuation rate of depth of thinking	0.69	Our survey evidence
Households			
β	Discount factor	0.99	$r = 4\%$ p.a.
σ^{-1}	Elasticity of intertemporal substitution (EIS)	1	
ν^{-1}	Frisch elasticity of labor supply	0.5	
Firms			
θ	Price stickiness	0.75	Average price duration of 1 year
κ	Phillips curve slope	0.086	$\kappa = \theta^{-1}(1-\theta)(1-\beta\theta)$
μ	Steady state markup	1.1	
Central bank			
ϕ	Taylor rule coefficient	1.5	

We follow a standard calibration of the New Keynesian model at quarterly frequency, with all parameters listed in Table 4.

4.1 Inflation Expectations and Long-Term Interest Rates

We study belief over- and underreaction in response to news about the cost-push shock ϵ_1^π and the monetary policy shock ϵ_1^i, as well as their consequences for period-0 long-term interest rates. We analyze the two shocks in sequence and present a synthesis afterward.

We assume that the yield of a 2-period bond $y_0^{(2)}$, i.e., the long-term yield, is determined by the expectations hypothesis as[8]

$$y_0^{(2)} = \frac{i_0 + \overline{\mathbb{E}}_0\left[i_1\right]}{2} \tag{19}$$

Cost-push shocks. Proposition 3 characterizes the period-1 equilibrium in response to the cost-push shock ϵ_1^π and beliefs thereof.

Proposition 3. (Period-1 cost-push shock) *The equilibrium response to a cost-push shock ϵ_1^π features*

$$\pi_1 = \frac{1}{1 - \underbrace{(1-\theta)}_{\text{pricing complementarity}} + \underbrace{\phi\theta\kappa}_{\text{monetary policy loop}} \cdot \underbrace{\sigma^{-1}(\nu+\sigma)}_{\text{Keynesian cross}}}\epsilon_1^\pi \tag{20}$$

$$i_1 = \phi\pi_1 = \frac{\phi}{1-(1-\theta)+\phi\theta\kappa\sigma^{-1}(\nu+\sigma)}\epsilon_1^\pi \tag{21}$$

whereas the period-0 average expectations, upon observing the news about ϵ_1^π, are

$$\overline{\mathbb{E}}_0\left[\pi_1\right] = \frac{1}{1-\lambda(1-\theta)+\lambda^4\phi\theta\kappa K_1(\lambda)}\epsilon_1^\pi \tag{22}$$

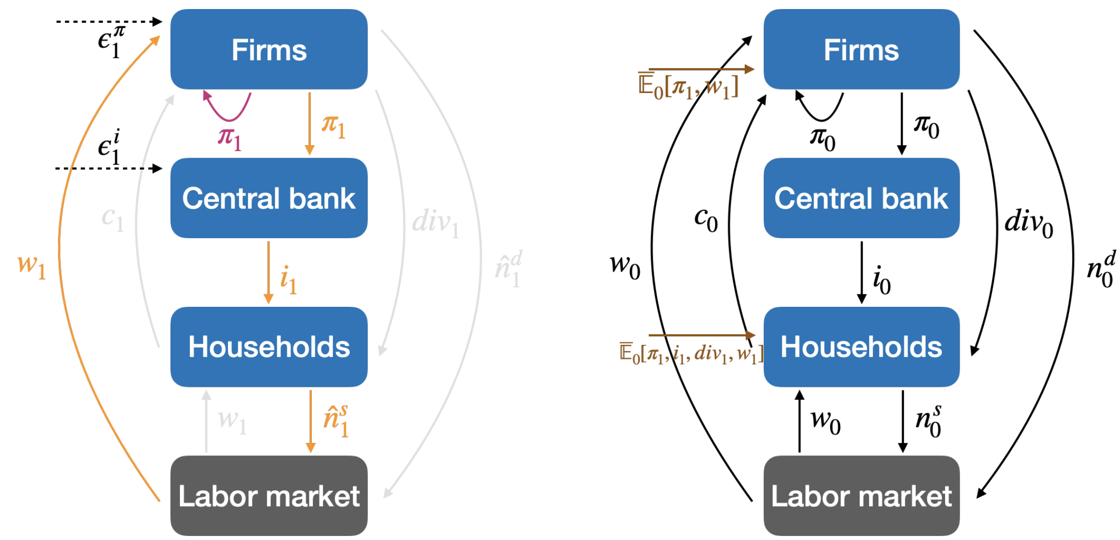
$$\overline{\mathbb{E}}_0\left[i_1\right] = \lambda\phi\overline{\mathbb{E}}_0\left[\pi_1\right] = \frac{\lambda\phi}{1-\lambda(1-\theta)+\lambda^4\phi\theta\kappa K_1(\lambda)}\epsilon_1^\pi \tag{23}$$

with $K_1(\lambda)$ increasing in λ and $K_1(1) = \sigma^{-1}(\nu+\sigma)$.

To understand these results, starting with inflation (20), recall that the equilibrium response is the sum of all n-step effects, as established in (9). We could organize all these effects into three groups that involve different loops, colored distinctively in Figure 3a, and the inflation response π_1 is directly involved in two of them.

The first loop is a self-loop of *pricing complementarity*, in purple. A higher inflation π_1 incentivizes all firms to price higher, thus amplifying itself. This effect has a strength of $(1-\theta)$. As we sum up the infinite series going through this loop, its strength appears in the denominator of (20).

[8]To microfound this in our model without aggregate risks, we assume that there is an intermediary who prices the 2-period bond on behalf of all households by averaging their beliefs.



(a) Causal relations of period-1 economy (b) Period-0 equilibrium

Figure 3: New Keynesian economy with transitory news shocks

The second one is a 4-step loop that involves the *monetary policy*, in orange. As inflation π_1 gets higher, the central bank raises the interest rate i_1, which encourages labor supply $\hat{n}_1^s$, which then lowers the real wage w_1. As a lower wage prompts the firms to lower their prices, this offsets the inflation response via 4 steps.

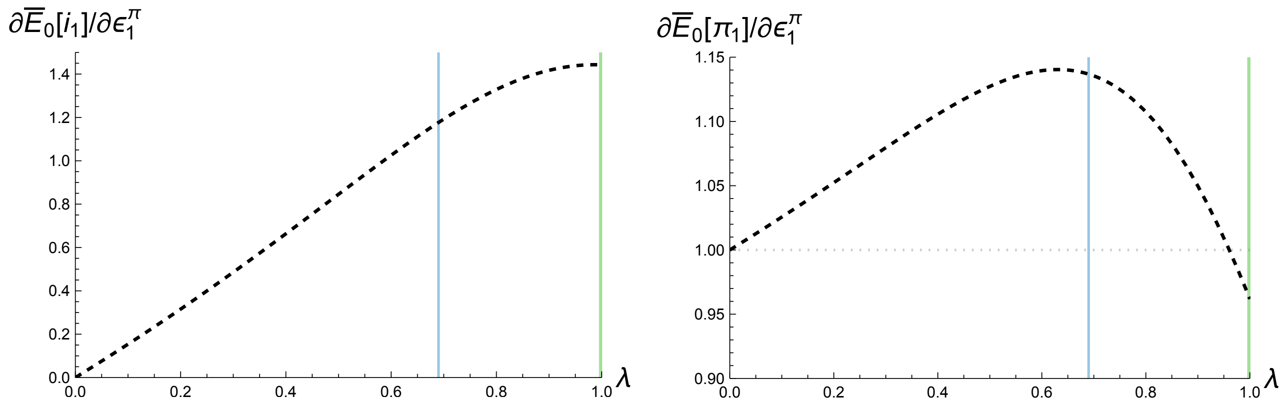
The final loop in the economy concerns the *Keynesian cross*, shown in gray. As the central bank raises the interest rate i_1, it discourages household consumption c_1, which leads the firm to pay lower dividends div_1 and reduce labor demand $\hat{n}^d$, resulting in a lower wage w_1. Consequently, households want to consume even less, triggering further adjustments of firms. This Keynesian cross strengthens any effect that impacts households, thereby compounding the monetary policy loop. Once again, summing the infinite geometric series results in the strength of this loop appearing in the denominator of (20).

The interest rate response i_1 (21) is simply the Taylor rule coefficient ϕ times inflation.

Turning to average beliefs, which are also sums of all n-step effects, but with deeper effects dampened more, as in (13). A loop that takes n steps to close will only be perceived by λ^n share of people once. As we add up the infinite series, the average inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$ (22) is modified relative to the true inflation (20) with different loops dampened to varying degrees. The longer 4-step loop is dampened more than the self-loop. In the limit of $\lambda = 0$, shallow agents do not perceive any propagation and we simply have $\overline{\mathbb{E}}_0[\pi_1] = \epsilon_1^\pi$.

The average interest expectation $\overline{\mathbb{E}}_0[i_1]$ (23) is λ times the Taylor rule coefficient ϕ times the inflation expectation, since it takes one more step for agents to appreciate the response of interest rate to inflation. In the limit of $\lambda = 0$, shallow agents do not perceive any response in the interest rate, i.e., $\overline{\mathbb{E}}_0[i_1] = 0$, nor do they perceive any variable other than inflation.

Figure 4 plots the interest rate expectation $\overline{\mathbb{E}}_0[i_1]$ and inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$ as functions of λ in dashed black lines, in response to cost-push shock ϵ_1^π. In each graph, the blue vertical line indicates our calibrated λ, whereas the green vertical link corresponds to the rational expectations, which coincide with the true responses i_1, π_1.



(a) Interest rate expectation $\overline{\mathbb{E}}_0[i_1]$

(b) Inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$

Figure 4: Average beliefs in response to period-1 cost-push shocks ϵ_1^π

Panel (a) implies that $\overline{\mathbb{E}}_0[i_1]$ underreacts to the cost-push shock compared to the true response, because agents underappreciate the Taylor rule. In our economy, the interest rate expectation $\overline{\mathbb{E}}_0[i_1]$ is a forward rate and a component of the long-term yield $y_0^{(2)}$. Thus our theory implies that the long-term yield itself will underreact. This is in line with findings in Bauer, Pflueger and Sunderam (2024) that surprises in the core consumer price index (CPI) led to very little changes in long-term yields until March 2022, when the Fed actually raised the federal funds rate.

Panel (b) suggests that the inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$ is non-monotonic in λ. Further, our calibration suggests that the inflation expectation exceeds the size of the direct effect of one, whereas the true inflation response is below one. That is, shallow agents think that the cost-push shock will be amplified, though actually it will be dampened. The underlying reason is that, in determining inflation (20), there is a shorter loop that amplifies the cost-push shock and a much longer loop that offsets it. When agents are shallow, they understand the shorter loop relatively better than the longer loop. Thus, on net, they perceive amplification. That can be true even if the longer offset loop is actually stronger than the shorter amplification loop, leading to actual net offset.

In the case with the responses of inflation π_1 and inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$ to the cost-push shock, shallow thinking flips the sign of the perceived net general equilibrium (GE) effect from offset to amplification. This occurs because the strong, long offsetting loop is severely dampened by shallow thinkers. The order of operations is key, like the Jensen's inequality. It would be a mistake to first collapse multiple GE effects in a model into a single net effect (which is offset in this case) and then naively dampen that effect. Indeed, when we write down the New Keynesian model, we carefully preserve the structural form of the equations as agents' best responses to decision-relevant variables, without substituting variables using equilibrium relations. For example, households respond to the wage w_1 and dividend div_1. If we had imposed that households respond to their total income, which equals c_1 in equilibrium, and introduced shallow thinking based on that equation, we would be assuming that agents understand that total income equals c_1.

Monetary policy shocks. Proposition 4 characterizes the period-1 equilibrium in response to the monetary policy shock ϵ_1^i and beliefs thereof.

Proposition 4. (Period-1 monetary policy shock) *The equilibrium response to a monetary policy shock ϵ_1^π features*

$$i_1 = \frac{1}{1 + \frac{\phi\theta\kappa\sigma^{-1}(\nu+\sigma)}{1-(1-\theta)}}\epsilon_1^i \tag{24}$$

$$\pi_1 = -\frac{\theta\kappa\sigma^{-1}(\nu+\sigma)}{1-(1-\theta)}i_1 = -\frac{\frac{\theta\kappa\sigma^{-1}(\nu+\sigma)}{1-(1-\theta)}}{1 + \frac{\phi\theta\kappa\sigma^{-1}(\nu+\sigma)}{1-(1-\theta)}}\epsilon_1^i \tag{25}$$

whereas the period-0 average expectations, upon observing the news about ϵ_1^i, are

$$\overline{\mathbb{E}}_0[i_1] = \frac{1}{1 + \frac{\lambda^4\phi\theta\kappa K_1(\lambda)}{1-\lambda(1-\theta)}}\epsilon_1^i \tag{26}$$

$$\overline{\mathbb{E}}_0[\pi_1] = -\frac{\lambda^3\theta\kappa K_1(\lambda)}{1-\lambda(1-\theta)}\overline{\mathbb{E}}_0[i_1] = -\frac{\frac{\lambda^3\theta\kappa K_1(\lambda)}{1-\lambda(1-\theta)}}{1 + \frac{\lambda^4\phi\theta\kappa K_1(\lambda)}{1-\lambda(1-\theta)}}\epsilon_1^i \tag{27}$$

with $K_1(\lambda)$ increasing in λ and $K_1(1) = \sigma^{-1}(\nu+\sigma)$.

These results relate to results in Proposition 3 regarding the cost-push shocks, but with a subtle and consequential difference. Starting with the interest rate (24), which is the sum of all n-step effects as in (9). All these effects belong to three different loops—pricing complementarity, the monetary policy loop, and the Keynesian cross—as established previously and displayed in Figure 3a.

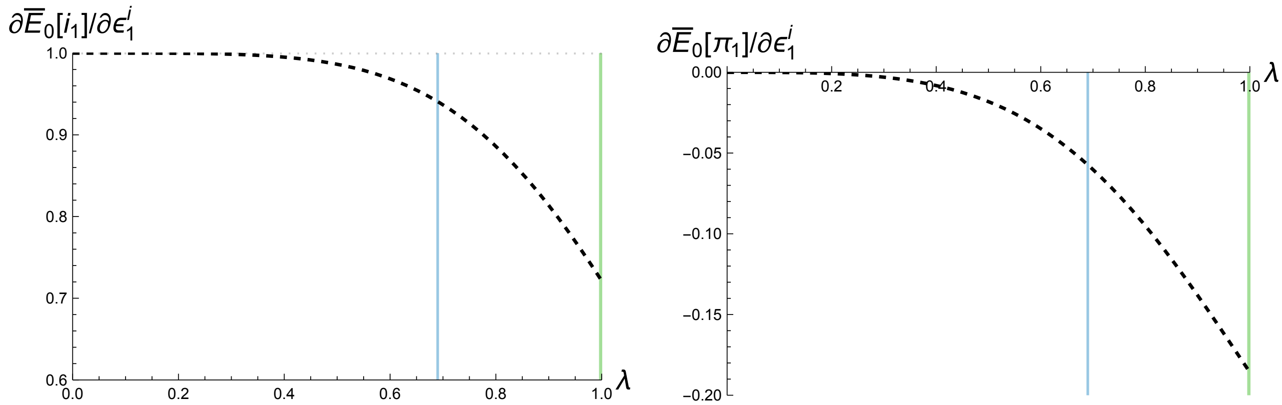
Among the three loops, the interest rate response is directly involved in *only one* of them, the monetary policy loop. And this loop offsets the interest rate response to a monetary policy shock via 4 steps: a higher interest rate i_1 encourages labor supply $\hat{n}_1^s$, which then lowers the real wage w_1, leading to less inflation π_1 through firms' pricing decisions, ultimately triggering the central bank to lower the interest rate i_1 according to the Taylor rule. This 4-step monetary policy offset loop with a strength $\phi\theta\kappa$ compounds with the other two loops, pricing complementarity and the Keynesian cross, which correspond to the $\frac{1}{1-(1-\theta)}$ and $\sigma^{-1}(\nu+\sigma)$ terms in (24). That is because pricing complementarity strengthens any effect on inflation and the Keynesian cross adds on any effect impacting households. Thus we conclude that *all* general equilibrium effects offset the interest response to a monetary policy shock, differing from the inflation response to a cost-push shock analyzed previously which involves both amplification and offset.

The inflation response π_1 (25) depends on the interest rate response i_1. It is again compounded by the pricing complementarity $\frac{1}{1-(1-\theta)}$ and the Keynesian cross $\sigma^{-1}(\nu+\sigma)$, since any effect of the interest rate on inflation goes through both firms and households.

Turning to average beliefs, which dampen deeper effects by more. In determining the interest rate expectation $\overline{\mathbb{E}}_0[i_1]$ (26), the 4-step loop is dampened by λ^4, pricing complementarity dampened by λ, and the Keynesian cross dampened as well captured by $K_1(\lambda)$, as in the case with inflation expectation (20) in response to cost-push shock. The inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$ derives accordingly from the interest rate expectation $\overline{\mathbb{E}}_0[i_1]$, but dampened by λ^3, since it takes 3 steps for the interest rate to affect the inflation. In the limit of $\lambda = 0$, shallow agents perceive no general equilibrium effects and thus $\overline{\mathbb{E}}_0[i_1] = \epsilon_1^i, \overline{\mathbb{E}}_0[\pi_1] = 0$.

Figure 5 plots the interest rate expectation $\overline{\mathbb{E}}_0[i_1]$ and inflation expectation $\overline{\mathbb{E}}_0[\pi_1]$ as functions of λ in dashed black lines, in response to monetary policy shock ϵ_1^i. As before, the blue vertical line indicates our calibrated λ, whereas the green vertical link corresponds to the rational expectations as well as the true responses i_1, π_1.

Panel (a) indicates that $\overline{\mathbb{E}}_0[i_1]$ overreacts to the monetary policy shock compared to the true response, thus suggesting that the period-0 long-term yield $y_0^{(2)}$ overreacts too.



(a) Interest rate expectation $\overline{\mathbb{E}}_0\left[i_1\right]$ (b) Inflation expectation $\overline{\mathbb{E}}_0\left[\pi_1\right]$

Figure 5: Average beliefs in response to period-1 monetary policy shocks ϵ_1^i

This aligns with a large literature on the excess sensitivity of long-term interest rates to monetary policy shocks (Cochrane and Piazzesi, 2002; Gürkaynak, Sack and Swanson, 2005; Hanson and Stein, 2015). The underlying mechanism is that, all general equilibrium effects offset the interest rate response to a monetary policy shock, and as they are dampened to varying degrees, shallow agents perceive less offset than the true overall effect. In contrast, panel (b) implies that $\overline{\mathbb{E}}_0\left[\pi_1\right]$ underreacts, because agents underappreciate the effect of monetary policy on inflation.

So far, we establish a rich pattern of conditional responses that the long-term interest rate overreacts to monetary policy shocks and underreacts to cost-push shocks, which has further implications for unconditional responses. With multiple shocks, if one runs a univariate regression of the long-term interest rate on the short-term rate, whether it suggests over- or underreaction depends on the mix of shocks. Indeed, Hanson, Lucca and Wright (2021) show that long-term interest rates are overly sensitive to changes in the short-term rates and predictably revert post 2000, suggesting overreaction. But the same pattern does not hold prior to 2000. Through the lens of our theory, one possibility is that there were more inflation shocks (like oil shocks) before 2000, contributing underreaction to the mix and obfuscating overreaction to monetary policy shocks. With multiple shocks, it is natural that one looks at a multivariate regression, which we visit in the next subsection.

Synthesis: strength and depth of GE effects jointly determine belief misreaction. To conclude this subsection, we offer a synthesis of belief misreaction. First consider a model with only one general equilibrium (GE) effect in addition to the partial equilibrium (PE) effect. In that model, if the GE effect amplifies (or offsets) the PE effect, beliefs underreact (or overreact) to shocks, as shallow agents underappreciate the GE effect.

However, our workhorse macroeconomic models feature multiple GE effects, in which case *both* the strength and depth of these GE effects matter for belief misreaction, as suggested by our theory. To understand this point, consider the inflation response to a cost-push shock ϵ_1^π. Under our calibration, shallow agents think that the cost-push shock ϵ_1^π will be amplified, even though it actually will be offset, i.e., $\frac{\partial \overline{\mathbb{E}}_0[\pi_1]}{\partial \epsilon_1^\pi} > 1 > \frac{\partial \pi_1}{\partial \epsilon_1^\pi}$. That occurs because shallow agents underappreciate the long, strong offset loop involving the monetary policy reaction, which is stronger when the Taylor rule coefficient ϕ is larger. We consider alternative values of ϕ in Figure 6.

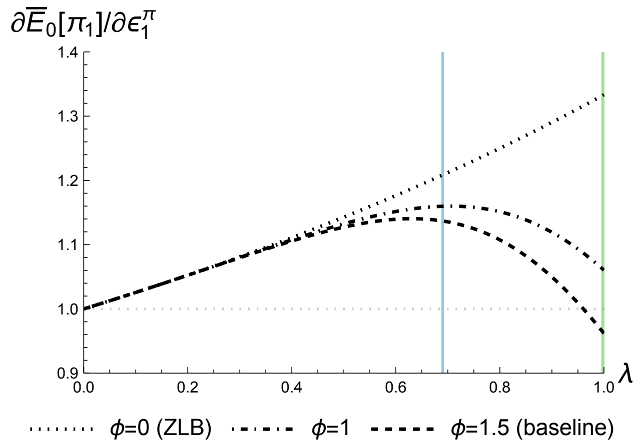


Figure 6: Inflation expectation $\overline{\mathbb{E}}_0\left[\pi_1\right]$ in response to cost-push shock ϵ_1^π, under various ϕ

If, instead, ϕ is of an intermediate value, e.g., $\phi = 1$, the long offset loop is too weak to turn the true net effect into offset. In that case, the true net effect is amplification, but shallow agents perceive even *more* amplification, i.e., $\frac{\partial \overline{\mathbb{E}}_0[\pi_1]}{\partial \epsilon_1^\pi} > \frac{\partial \pi_1}{\partial \epsilon_1^\pi} > 1$, as they overweigh the short amplification loop relative to the long offset loop.

With an even lower ϕ, for example, if the policy rate is stuck at the zero lower bound (ZLB) with $\phi = 0$,[9] the long offset loop is weak or non-existent, leaving the short amplification loop as the dominant force. In that case, we are almost in a model with only one GE

[9]The Taylor principle requires $\phi > 1$ for determinacy of the infinite-horizon New Keynesian model under rational expectations. In our 2-period setting, determinacy is not a concern. With persistent shocks, shallow thinking beliefs are uniquely defined by a formula similar to (10), which helps select an equilibrium. For further discussion on determinacy, which is not the focus of our paper, see Farhi and Werning (2019), Gabaix (2020) and Angeletos and Lian (2023*b*).

effect, and it amplifies the PE effect. As a result, shallow agents perceive less amplification on net than the true effect, i.e., $\frac{\partial \pi_1}{\partial \epsilon_1^\pi} > \frac{\partial \overline{\mathbb{E}}_0[\pi_1]}{\partial \epsilon_1^\pi} > 1$.

4.2 Predictability of Bond Excess Returns

We show shallow thinking implies that bond excess returns can be predicted by macroeconomic variables other than the current yields, as established empirically by Joslin, Priebsch and Singleton (2014), Cieslak (2018) and others.

We consider the excess return of holding a 2-period (i.e., long-term) bond from period 0 to period 1, relative to holding a 1-period (i.e., short-term) bond,

$$xr_{0\to 1}^{(2)} \equiv \underbrace{-i_1 + 2y_0^{(2)}}_{\text{return of long-term bond}} - \underbrace{i_0}_{\text{return of short-term bond}} = \overline{\mathbb{E}}_0 [i_1] - i_1 \tag{28}$$

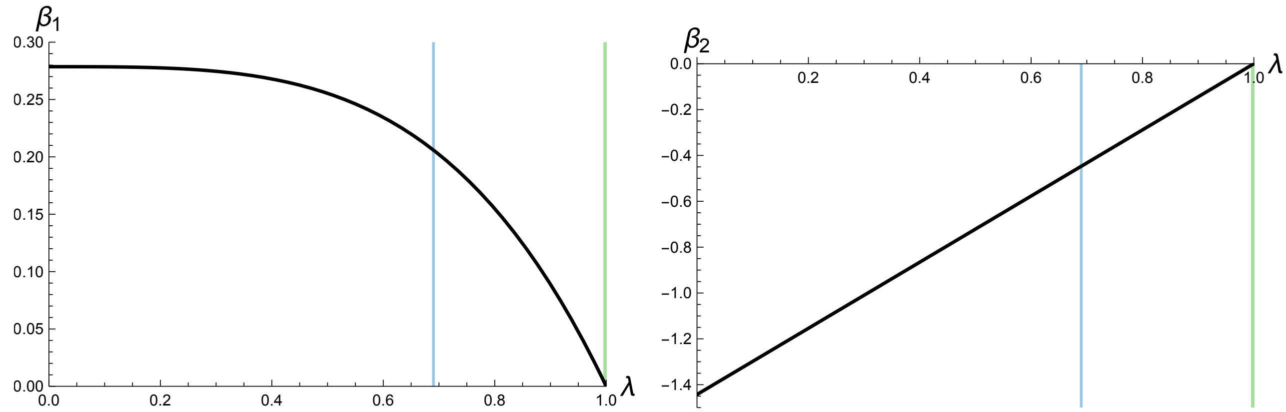
where the equality follows from (19). The intuition is very simple, when the period-0 expectation of period-1 interest rate exceeds its actual value, the long-term bond is undervalued in period 0 and will appreciate in period 1, leading to a positive excess return, and vice versa.

We study its predictability by the inflation expectation $\overline{\mathbb{E}}_0 [\pi_1]$, as studied by Joslin, Priebsch and Singleton (2014) and Cieslak (2018), while controlling for the forward rate $\overline{\mathbb{E}}_0 [i_1]$

$$\underbrace{xr_{0\to 1}^{(2)}}_{\text{bond excess return}} = \beta_1 \underbrace{\overline{\mathbb{E}}_0 [i_1]}_{\text{forward rate}} + \beta_2 \underbrace{\overline{\mathbb{E}}_0 [\pi_1]}_{\text{inflation expectation}} + \alpha + \epsilon_{0\to 1} \tag{29}$$

Figure 7 illustrates the theory-implied coefficients β_1 and β_2 as functions of λ in solid black lines. As long as $\lambda < 1$, our theory leads to a negative β_2, which is what Joslin, Priebsch and Singleton (2014) and Cieslak (2018) found when using inflation expectations or some other macroeconomic variables to predict bond excess returns.

To understand the mechanism, first note that in the limit of rational expectations ($\lambda = 1$), there is no excess return, hence $\beta_1, \beta_2 = 0$. In the other limit of $\lambda = 0$, we have $\overline{\mathbb{E}}_0 [i_1] = \epsilon_1^i, \overline{\mathbb{E}}_0 [\pi_1] = \epsilon_1^\pi$ from Propositions 3 and 4. This case implies $\beta_1 > 0$ and $\beta_2 < 0$, because a monetary policy shock leads to overreaction of interest expectation and hence a positive excess return, while a cost-push shock leads to underreaction of interest rate expectation and thus a negative excess return. With any intermediate value of λ, we still have $\beta_1 > 0$ and $\beta_2 < 0$, as the forward rate primarily reflects the monetary policy shock,



(a) Coefficient β_1 of forward rate $\overline{\mathbb{E}}_0 [i_1]$ (b) Coefficient β_2 of inflation expectation $\overline{\mathbb{E}}_0 [\pi_1]$

Figure 7: Predictability of bond excess returns by macroeconomic variables

and the inflation expectation mostly loads on the cost-push shock.

Taking stock of findings here and in the previous subsection, we conclude that shallow thinking offers a unified rationalization of several bond market puzzles that seem unrelated or even contradictory. These include the underreaction of long-term interest rates (Bauer, Pflueger and Sunderam, 2024), their overreaction (Cochrane and Piazzesi, 2002; Gürkaynak, Sack and Swanson, 2005; Hanson and Stein, 2015), as well as the predictability of excess bond returns (Joslin, Priebsch and Singleton, 2014; Cieslak, 2018).

4.3 Effects of Cost-Push Shocks

Now we establish the consequences of shallow thinking regarding aggregate outcomes in period 0, in response to the transitory news shocks. In particular, we show that news about period-1 cost-push shocks are more inflationary and less contractionary than the rational expectations prediction.

Period-0 equilibrium. In period 0, the shocks have not materialized but firms and households are forward-looking. The period-0 equilibrium consists of seven variables $\left\{i_0, \pi_0, div_0, n_0^d, c_0, n_0^s, w_0\right\}$, similar to the period-1 equilibrium. Among these seven variables, three of them (the interest rate i_0, dividend div_0 and labor demand n_0^d) depend only on the contemporaneous values of the other variables. These contemporaneous causal relations are the same as their period-1 counterparts (1, 2, 6). Three variables (the inflation π_0, consumption c_0 and labor supply n_0^s) depend on agents' beliefs about period-1 outcomes, since firms's pricing decisions and households' consumption and labor supply decisions are forward-looking, detailed next. Last, the wage w_0 arises from the labor market clearing condition $n_0^s = n_0^d$.

The period-0 inflation π_0 satisfies

$$\pi_0 = \theta\kappa\left(w_0 + \beta\theta\overline{\mathbb{E}}_0\left[w_1\right]\right) + (1-\theta)\left(\pi_0 + \beta\theta\overline{\mathbb{E}}_0\left[\pi_1\right]\right) \tag{30}$$

which increases in expectations of both future real wage w_1 and inflation π_1, as firms want to front run a higher future marginal cost.

The period-0 consumption c_0 and labor supply n_0^s follow

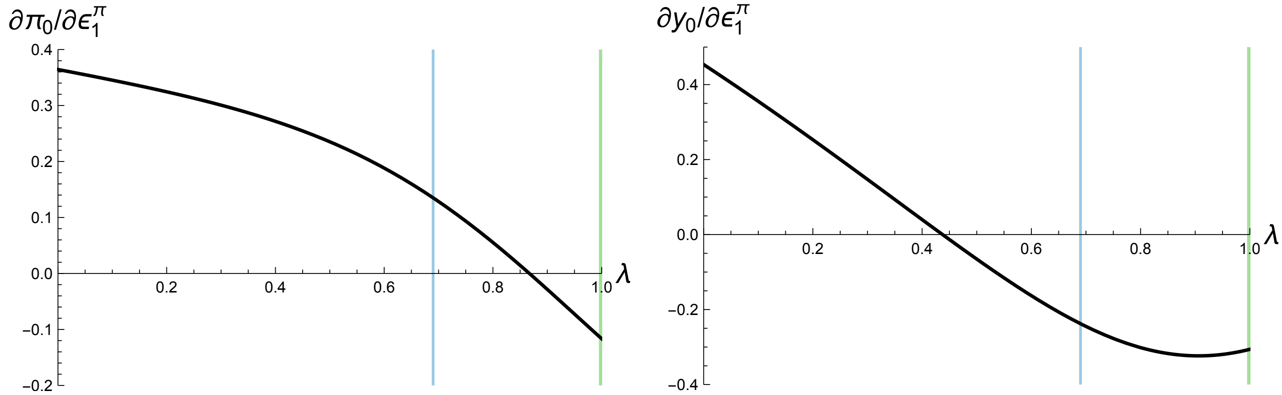
$$\begin{aligned} c_0 = &-\sigma^{-1}\beta\left(i_0 - \overline{\mathbb{E}}_0\left[\pi_1\right] + \beta\overline{\mathbb{E}}_0\left[i_1\right]\right) + \frac{(1-\beta)(\mu-1)\nu}{\sigma+\mu\nu}\left(div_0 + \overline{\mathbb{E}}_0\left[div_1\right]\right) \\ &+ \frac{(1-\beta)(1+\nu)}{\sigma+\mu\nu}\left(w_0 + \overline{\mathbb{E}}_0\left[w_1\right]\right) \end{aligned} \tag{31}$$

$$\begin{aligned} n_0^s = &\,\nu^{-1}\beta\left(i_0 - \overline{\mathbb{E}}_0\left[\pi_1\right] + \beta\overline{\mathbb{E}}_0\left[i_1\right]\right) - \frac{(1-\beta)(\mu-1)\sigma}{\sigma+\mu\nu}\left(div_0 + \overline{\mathbb{E}}_0\left[div_1\right]\right) \\ &+ \nu^{-1}\left[1 - \sigma\frac{(1-\beta)(1+\nu)}{\sigma+\mu\nu}\right]w_0 - \frac{(1-\beta)(1+\nu)\nu^{-1}\sigma}{\sigma+\mu\nu}\overline{\mathbb{E}}_0\left[w_1\right] \end{aligned} \tag{32}$$

Households react to the future interest rate, dividend, and wage, as well as the future inflation, as a higher inflation lowers the real interest rate from period 0 to period 1.

In determining the period-0 equilibrium, the only decisions-relevant beliefs are about inflation π_1, wage w_1, interest rate i_1 and dividend div_1. Figure 3b illustrates causal system of period-0 equilibrium, where these beliefs act like shocks to firms and households. In particular, the inflation expectation $\overline{\mathbb{E}}_0\left[\pi_1\right]$ serves two roles here: it acts like a cost-push shock for firms as they want to front run future inflation, and it functions as a demand shock for households since a higher inflation lower the real interest rate. This observation will prove useful when we discuss the effects of cost-push shocks next.

Cost-push shocks. Figure 8 illustrates the period-0 equilibrium, in response to period-1 cost-push shock ϵ_1^π, in solid black lines. The period-0 output is simply the consumption c_0. Still the blue line indicates our calibration of shallow thinking λ whereas the green line stands for rational expectations. Unlike the period-1 equilibrium, which is independent of λ, the period-0 equilibrium does depend on λ. Thus the green line represents the rational expectations equilibrium (REE).



(a) Period-0 inflation π_0

(b) Period-0 output y_0 (which equals consumption c_0)

Figure 8: Period-0 equilibrium in response to period-1 cost-push shocks ϵ_1^π

Panel (a) suggests that a period-1 cost-push shock is *inflationary* in period 0 under shallow thinking, but *deflationary* under rational expectations. It is deflationary under rational expectations because rational agents anticipate that the central bank will raise interest rate i_1 in response to higher inflation π_1, leading to a contraction of the period-1 economy. As a result, households cut back on their consumption today, causing a contract in period 0 as well, as seen in panel (b). In contrast, in the limit of extremely shallow agents ($\lambda = 0$), they only perceive a change in future inflation, $\overline{\mathbb{E}}_0[\pi_1] = \epsilon_1^\pi$, but not in any other variable. As we have established with (30-32), inflation expectations act like a cost-push shock for firms and a demand shock for households. As firms price higher and households consume more and work less, this leads to period-0 inflation and output expansion. With an intermediate value of λ, agents somewhat expect a future bust, but not fully. Under our estimated value of λ, period-1 cost-push shocks lead to inflation and a smaller output contraction in period 0.

From transitory news shocks to persistent shocks. So far, we have focused on transitory news shocks, but this is much more informative than it may seem. As we have noted before, comparing theories of beliefs under transitory news shocks is equivalent to comparing them under two-period shocks, since the period-0 equilibrium response to contemporaneous shocks is independent of belief formation. Furthermore, analyzing transitory news shock can help build intuition for persistent shocks. A persistent cost-push shock observed in period 0 is technically a sum of a current shock and a series of future shocks. Shallow thinking implies that a future cost-push shock is inflationary in period 0, whereas rational expectations predict otherwise. This suggests that a persistent cost-push shock is more inflationary under shallow thinking than under rational expectations. Additionally, a *more* persistent cost-push shock is simply a sum of more future shocks. Thus, shallow thinking predicts that a *more* persistent cost-push shock leads to *higher* inflation, contrary to the rational expectations prediction.

In the next subsection, we will generalize shallow thinking to persistent shocks to formalize these two findings. That generalization will fill in one missing piece from the intuitive analysis above: in responding to a persistent shock, agents must think about how the future period-t equilibrium depends on shocks in further periods s.

5 Consequences of Persistent Shocks

We generalize shallow thinking to accommodate persistent shocks in the New Keynesian economy and show that the insights gained from transitory news shocks still hold, while new lessons emerge. We introduce shallow thinking of a dynamic general equilibrium in Section 5.1, by focusing on the cross-variable causal relations and abstracting away from the intertemporal dimension. This will allow us to formally address belief misreaction to persistent shocks and their effects in Section 5.2. Readers who are more interested in the consequences and less interested in the technical details could skip the first subsection.

We consider persistent shocks $\{\epsilon_t\}_{t\geq 0}$ that are observed at time 0^-, and assume that agents form expectations once and for all, i.e., $\overline{\mathbb{E}}_t[\cdot] \equiv \overline{\mathbb{E}}[\cdot]$. This assumption simplifies the analysis, as it nests the rational expectations benchmark and is reasonable for our theory for the following two reasons.

First, an important reason for expectation updates over time is that agents gradually learn about shocks. We focus on the rationality of beliefs in the absence of any information frictions, and our survey design mimics this environment.

Second, expectations can update as agents learn how aggregate variables respond to shocks through repeated experiences, similar to how economists study shock propagation using time series data. We do not explicitly model such learning processes, but we note that in measuring shallow thinking through a survey, we consider conventional shocks, such as oil shocks and monetary policy shocks. These are age-old shocks, so our estimation already incorporates knowledge that people have gained over time. This contrasts with unconventional policy shocks (such as forward guidance or quantitative easing), to which people had no prior exposure until recently. That said, how agents form their mental models through learning and how this interacts with information frictions are promising topics for future research.

5.1 Shallow Thinking of Dynamic General Equilibrium

The infinite-horizon New Keynesian model is characterized by sequences of seven variables $\left\{i_t, \pi_t, div_t, n_t^d, c_t, n_t^s, w_t\right\}_{t\geq 0}$. We will refer to this collection of variables as $\mathcal{V}$. The first six variables are agents' actions, which we collect as $\mathcal{V}^{action}$, whereas the last is a price formed in the competitive labor market. Note that we start with the labor supply and demand n_t^s, n_t^d in order to reinterpret them as supply and demand curves shortly. In that reinterpretation, with slight abuse of notation, we will use $\mathcal{V}$ and $\mathcal{V}^{action}$ to denote variables of interests with n^s, n^d replaced by $\hat{n}^s, \hat{n}^d$.

With persistent shocks, we distinguish between two sets of equations that jointly determine the equilibrium outcomes, similar to García-Schmidt and Woodford (2019). The first set corresponds to relations among economic variables that arise from optimal decisions of economic agents, given the current realizations and their expected future values of variables that directly affect them. We call these temporary equilibrium relations. Based on these temporary equilibrium relations, we will define the second set of equations that characterizes expectations, which generalizes Assumption 1.

In terms of the temporary equilibrium relations, three variables i_t, div_t, n_t^d only depend on contemporaneous values of other variables, laid out in (A1, A2, A6). Three variables π_t, c_t, n_t^s are forward-looking. Inflation π_t is a linear function in w_t, π_t and $\left\{\overline{\mathbb{E}}_t\left[w_{t+k}\right], \overline{\mathbb{E}}_t\left[\pi_{t+k}\right]\right\}_{k\geq 1}$, and consumption and labor supply c_t, n_t^s are linear functions in i_t, div_t, w_t and $\left\{\overline{\mathbb{E}}_t\left[i_{t+k}\right], \overline{\mathbb{E}}_t\left[\pi_{t+k}\right], \overline{\mathbb{E}}_t\left[div_{t+k}\right], \overline{\mathbb{E}}_t\left[w_{t+k}\right]\right\}_{k\geq 1}$, detailed in (A3-A5). The last variable w_t arises from equilibrating labor supply n_t^s and demand n_t^d. These seven equations completely characterize the equilibrium, given beliefs.

To determine shallow thinking beliefs, we start by characterizing the rational expectations equilibrium (REE) and causal relations thereof. By replacing each expectation $\overline{\mathbb{E}}_t\left[v_\tau\right]$ with the true outcome v_τ under rational expectations, the six variables in $\mathcal{V}^{action}$ that agents choose can be represented in the sequence space, following Farhi and Werning (2019) and Auclert, Rognlie and Straub (2024), as

$$\mathbf{v}^{REE} = \sum_{\mathbf{u}\in\mathcal{V}} \mathbf{J}_{\mathbf{vu}}\mathbf{u}^{REE} + \boldsymbol{\epsilon}^v,\ \forall \mathbf{v}\in\mathcal{V}^{action} \tag{33}$$

where

$$(\mathbf{J_{vu}})_{t\tau} \equiv \begin{cases} \frac{\partial v_t}{\partial u_\tau} & \tau \le t \\ \frac{\partial v_t}{\partial \overline{\mathbb{E}}_t[u_\tau]} & \tau > t \end{cases}$$

is the Jacobian of the sequence of one variable $\mathbf{v} \equiv (\{v_t\}_{t\ge 0})'$ with respect to the sequence of another variable $\mathbf{u} \equiv (\{u_t\}_{t\ge 0})'$, and $\boldsymbol{\epsilon}^v \equiv \left(\left\{\epsilon_t^v\right\}_{t\ge 0}\right)'$ denotes the sequence of a structural shock. The first term in the parenthesis captures best response to current and past realizations, and the second term embeds the best response to beliefs about future outcomes. The Jacobians $\mathbf{J_{vu}}$ are upper triangular matrices in this New Keynesian model which is purely forward-looking with no state variable, but the formulation (33) is general to accommodate models with state variables that depend on the past, such as the RBC model.

For labor supply and demand $\mathbf{n}^s, \mathbf{n}^d$, by separating their dependence on the wage $\mathbf{w}$ from the rest, we interpret them as supply and demand curves in the sequence space

$$\mathbf{n}^{s,REE} = \mathbf{J}_{\mathbf{n}^s\mathbf{w}}\mathbf{w}^{REE} + \hat{\mathbf{n}}^{s,REE}, \quad \mathbf{n}^{d,REE} = \mathbf{J}_{\mathbf{n}^d\mathbf{w}}\mathbf{w}^{REE} + \hat{\mathbf{n}}^{d,REE}$$

with elasticities $\mathbf{J}_{\mathbf{n}^s\mathbf{w}}, \mathbf{J}_{\mathbf{n}^d\mathbf{w}}$ and shifters $\hat{\mathbf{n}}^{s,REE}, \hat{\mathbf{n}}^{d,REE}$ defined as

$$\hat{\mathbf{n}}^{s,REE} = \sum_{\mathbf{u}\in\mathcal{V}^{action}} \mathbf{J}_{\mathbf{n}^s\mathbf{u}}\mathbf{u}^{REE} + \boldsymbol{\epsilon}^{n^s}, \quad \hat{\mathbf{n}}^{d,REE} = \sum_{\mathbf{u}\in\mathcal{V}^{action}} \mathbf{J}_{\mathbf{n}^d\mathbf{u}}\mathbf{u}^{REE} + \boldsymbol{\epsilon}^{n^d} \tag{34}$$

Demand elasticity $\mathbf{J}_{\mathbf{n}^d\mathbf{w}}$ is a matrix of zeros in this model since firms only use labor inputs, but, more generally (such as in the RBC model), it does not have to be. Supply elasticity $\mathbf{J}_{\mathbf{n}^s\mathbf{w}}$ is a upper triangular matrix, as households' labor supply responds to future wages. To specify beliefs, we reason with $\hat{\mathbf{n}}^{s,REE}, \hat{\mathbf{n}}^{d,REE}$ instead of $\mathbf{n}^{s,REE}, \mathbf{n}^{d,REE}$ in $\mathcal{V}$ and $\mathcal{V}^{action}$.

By equalizing $\mathbf{n}^{s,REE} = \mathbf{n}^{d,REE}$, we can interpret the wage $\mathbf{w}^{REE}$ as resulting from

$$\mathbf{w}^{REE} = (\mathbf{J}_{\mathbf{n}^s\mathbf{w}} - \mathbf{J}_{\mathbf{n}^d\mathbf{w}})^{-1}\left(\hat{\mathbf{n}}^{d,REE} - \hat{\mathbf{n}}^{s,REE}\right) \tag{35}$$

This rule of price determination generalizes its counterpart in the period-1 economy (7) and incorporates the response of time-t wage on time-s demand and supply shifts.

Taking stock of the rational expectations equilibrium, (33) describes all agents' actions $\mathcal{V}^{action}$ and (35) characterizes the wage from the competitive labor market. We stack the sequences of outcomes to form a long vector $\mathbf{V} \equiv (\{\mathbf{v}'\}_{v\in\mathcal{V}})'$, sequences of shocks as $\mathbf{S} \equiv \left(\left\{\boldsymbol{\epsilon}^{v'}\right\}_{v\in\mathcal{V}}\right)'$, and correspondingly stack the Jacobians as a giant matrix $\mathbf{M} \equiv \left(\{\mathbf{J_{vu}}\}_{v,u\in\mathcal{V}}\right)$.

The rational expectations equilibrium satisfies a *fixed point* system

$$\underbrace{\mathbf{V}^{REE}}_{\text{sequence of variables}} = \underbrace{\mathbf{M}}_{\text{sequence-space causal relations}} \mathbf{V}^{REE} + \underbrace{\mathbf{S}}_{\text{sequence of shocks}} \tag{36}$$

which can also be solved as a *sum* of all effects

$$\mathbf{V}^{REE} = (\mathbf{I} - \mathbf{M})^{-1}\,\mathbf{S} = \sum_{n=1}^{\infty} \mathbf{M}^{n-1}\mathbf{S} \tag{37}$$

where each $\mathbf{M}^{n-1}$ term is an n-step effect of the sequence of a shock on the sequence of a variable via $n-1$ variables. These generalize their period-1 counterparts (8, 9).

Assumption 1′. Individuals vary in their finite *depth of thinking* $d \in \mathbb{N}^+$,

$$\mathbb{E}^d\,[\mathbf{V}] \equiv \sum_{n=1}^{d} \mathbf{M}^{n-1}\mathbf{S} \tag{38}$$

Assumptions 1′ and 2 jointly lead to the following characterizations of the average beliefs, generalizing (13, 14) in Proposition 1,

$$\overline{\mathbb{E}}\,[\mathbf{V}] = \sum_{n=1}^{\infty} \lambda^{n-1}\mathbf{M}^{n-1}\mathbf{S} \tag{39}$$

$$= \lambda\mathbf{M}\overline{\mathbb{E}}\,[\mathbf{V}] + \mathbf{S} \tag{40}$$

Given the average beliefs $\overline{\mathbb{E}}\,[\cdot]$, the equilibrium $\left\{i_t, \pi_t, div_t, n_t^d, c_t, n_t^s, w_t\right\}_{t\geq 0}$ is characterized by temporary equilibrium relations (A1-A6) as well as the labor market clearing condition $n_t^s = n_t^d$.

5.2 Persistent Cost-Push Shocks

We consider responses of beliefs and equilibrium to a persistent cost-push shock $\epsilon_t^\pi = \rho^t \epsilon^\pi$. In response to such an exponentially decaying shock, beliefs and equilibrium outcomes are exponentially decaying at the same rate. Proposition 5 establishes the rational expectations equilibrium and shallow thinking beliefs.

Proposition 5. (Persistent cost-push shock) *The rational-expectations equilibrium (REE) re-*

sponse to a persistent cost-push shock $\epsilon_t^\pi = \rho^t \epsilon^\pi$ features

$$\pi_t^{REE} = \left[1 - \frac{(1-\theta) + \left(\rho - \phi\right)\theta\kappa\frac{\sigma^{-1}(\nu+\sigma)}{1-\rho}}{1-\beta\theta\rho}\right]^{-1} \epsilon_t^\pi \tag{41}$$

$$i_t^{REE} = \phi\pi_t^{REE} = \left[1 - \frac{(1-\theta) + \left(\rho - \phi\right)\theta\kappa\frac{\sigma^{-1}(\nu+\sigma)}{1-\rho}}{1-\beta\theta\rho}\right]^{-1} \phi\epsilon_t^\pi \tag{42}$$

whereas the average expectations under shallow thinking are

$$\overline{\mathbb{E}}\left[\pi_t\right] = \left[1 - \frac{\lambda(1-\theta) + \left(\lambda^3\rho - \lambda^4\phi\right)\theta\kappa K(\lambda,\rho)}{1-\beta\theta\rho}\right]^{-1} \epsilon_t^\pi \tag{43}$$

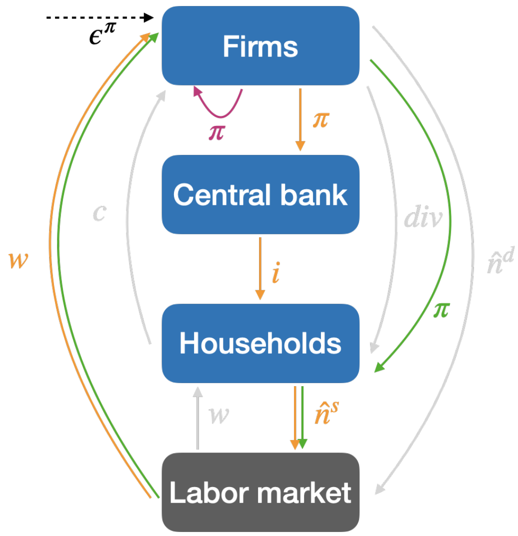
$$\overline{\mathbb{E}}\left[i_1\right] = \lambda\phi\overline{\mathbb{E}}_0\left[\pi_1\right] = \left[1 - \frac{\lambda(1-\theta) + \left(\lambda^3\rho - \lambda^4\phi\right)\theta\kappa K(\lambda,\rho)}{1-\beta\theta\rho}\right]^{-1} \lambda\phi\epsilon_t^\pi \tag{44}$$

with $K(\lambda,\rho)$ increasing in λ and ρ and $K(1,\rho) = \frac{\sigma^{-1}(\nu+\sigma)}{1-\rho}$.

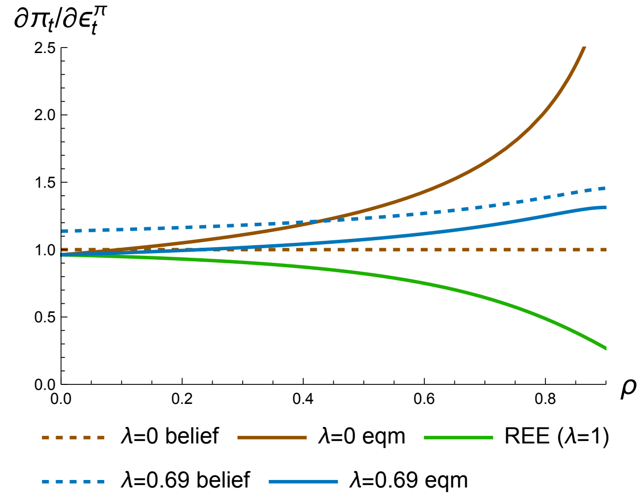
This proposition shows how exactly the persistence of shock ρ matters, nesting Proposition 3 with $\rho = 0$. A positive ρ gives rise to new terms and modifies existing terms, which we dissect in order by analyzing REE inflation (41) and shallow expectations (43).

Regarding new terms, four general equilibrium loops across sequences of variables are involved, instead of three, displayed in Figure 9a. Relative to the case with period-1 shocks (Figure 3a), persistent shocks give rise to new causal relations—the responses of household consumption and labor supply to inflation π, plotted in green. This results in the fourth loop, the *inflation-on-households loop* (in green), in addition to pricing complementarity (in purple), the monetary policy loop (in orange), and the Keynesian cross (in gray). This loop takes three steps to close and amplifies the inflation response to cost-push shocks, since a higher inflation π simultaneously encourages consumption c and discourages labor supply $\hat{n}^s$, which leads to a higher wage w, feeding into inflation π. As a result, this loop is dampened by λ^3 in expectations (43). Further, the strength of this loop is proportional to ρ, meaning that it only exits when $\rho > 0$ and is stronger when ρ is higher. That occurs because only future inflation impacts household behavior by changing the real interest rate, and the expected inflation decays with rate ρ.

Concerning the strength of these terms, we note that a positive ρ does two things, in



(a) Causal relations in the sequence space



(b) Inflation π_t

Figure 9: New Keynesian economy with persistent shocks

addition to activating the inflation-on-households loop. First, it strengths all loops by the amount of $\frac{1}{1-\beta\theta\rho}$, because firms' pricing decisions are forward-looking and respond to a sum of future changes discounted by $\beta\theta$. Second, it further boosts the Keynesian cross via $K(\lambda, \rho)$. This additional boost occurs because households' decisions are forward-looking as well. Since the Keynesian cross reflects the feedback between firms and households, the persistence of shocks is compounded. As ρ increases, all loops get stronger, but the inflation-on-households and monetary policy loops become even stronger relative to the pricing complementarity self-loop, which has important consequences, as we will discuss next.

Figure 9b plots the average inflation expectation $\overline{\mathbb{E}}[\pi_t]$ and the equilibrium inflation π_t, both relative to the time-t size of the shock, as functions of persistence ρ. The dashed lines indicate expectations and the solid lines represent the equilibrium. Different colors stand for different values of λ, with $\lambda = 0$ in brown, $\lambda = 0.69$ in blue (our calibration), and $\lambda = 1$ (rational expectations) in green. In the case with rational expectations, the beliefs coincide with the equilibrium.

We note from Figure 9b that two insights obtained with transitory news shocks extend here, and a new lesson emerges regarding persistence. First, inflation expectations exceed their equilibrium values under shallow thinking, as the dashed blue line lies above the solid blue line. Second, cost-push shocks are more inflationary under shallow thinking than under rational expectations, since the solid blue line lies above the solid green line. These two findings generalize the previous results from transitory news shocks. Last, a persistent cost-push shock leads to higher inflation under shallow thinking, but lower inflation under rational expectations, since the solid blue line increases in ρ while the solid green line decreases in ρ.

To understand this new lesson regarding persistence, we start with the limit case of $\lambda = 0$ and $\lambda = 1$ (rational expectations). In the limit of $\lambda = 0$, inflation expectations always have the same size as the shock, with the dashed brown being flat. Inflation expectations change agents' behavior by acting like a cost-push shock for firms and a demand-shock for households. As their decisions are forward-looking and depend on discounted sums of future disturbances, a more persistent shock leads to larger changes in their behavior in any period, thus resulting in higher inflation.

In contrast, in the limit of $\lambda = 1$ (rational expectations), as analyzed based on Proposition 5, a higher ρ strengthens all general equilibrium effects and, further, boosts the inflation-on-households and monetary policy loops relative to the pricing complementarity self-loop. The monetary policy loop offsets the inflation response, whereas the other two amplifies it. As the monetary policy loop is the strongest among them and rational agents appreciate that, a higher ρ strengthens the offset and leads to a lower inflation. That is, facing a more persistent shock, rational agents anticipate that the monetary policy reaction will offset it more by acting on the Keynesian cross between firms and households, and their beliefs coincide with the rational expectations equilibrium.

Under our calibrated value $\lambda = 0.69$, shallow agents' inflation expectations increase in ρ. The mechanism is that, while the monetary policy offset loop is objectively the strongest, it is also the longest and gets dampened more in expectations. Thus, shallow agents believe that a more persistent shock leads to more amplification, as they better understand the shorter amplification loops than the monetary policy offset loop. As a result, the equilibrium inflation is higher when ρ is higher. That increasing relationship is less drastic than in the extreme case of $\lambda = 0$, as shallow agents partially understand the monetary policy reaction and its pressure on the economy.

6 Consequences in an RBC Economy

6.1 Shallow Thinking in an RBC Economy

6.2 Effects of Productivity Shocks

7 Conclusion

This paper develops a theory of *shallow thinking* as the structure of belief formation, supports its empirical content using a customized survey, and illustrates its consequences for macroeconomics and finance.

The key implication of our theory is that deeper causal relations have less influence on beliefs, meaning both the strength and depth of these relations matter. Our estimation suggests that the average depth of thinking is only about three—far below infinity assumed by rational expectations. We also provide evidence that this limited depth of thinking reflects people's limited knowledge about the macroeconomy. While our primary contribution is to develop a psychologically grounded model of expectations for macroeconomic analysis, our study of perceived macroeconomic shock propagation also advances the causal reasoning literature (Waldmann, 2017). These works typically present participants with simple examples in short experiments, whereas we examine a real-world domain that involves many variables and long-term data accumulation.

Our theory leads to a host of consequences for macro and finance. In a New Keynesian model, long-term nominal rates overreact to monetary shocks and underreact to cost-push shocks, as agents underappreciate offsetting loops and the propagation of shocks. This insight reconciles multiple bond market puzzles. Additionally, cost-push shocks are more inflationary than predicted under rational expectations, and *more* persistent cost-push shocks lead to *higher* inflation—contrary to rational expectations' predictions. That occurs because shallow agents better understand shallow general equilibrium effects, which amplify cost-push shocks. In a real business cycle model, shallow thinking amplifies fluctuations in response to productivity shocks, in a way consistent with observed return predictability over investment boom-bust cycles.

Stepping back, we acknowledge the immense complexity of the economy. If it has taken decades for our best economists to understand how it works—or if they are still figuring it out—we must carefully consider how much the average person understands. Our theory of shallow thinking in general equilibrium may be fruitfully applied to many other settings, and we hope that this research agenda can make meaningful contributions to economics.

Bibliography

Afrouzi, Hassan, Spencer Y. Kwon, Augustin Landier, Yueran Ma, and David Thesmar. 2023. "Overreaction in Expectations: Evidence and Theory." *The Quarterly Journal of Economics*, 138(3): 1713–1764. 7

Ahn, Woo-kyoung, Nancy S. Kim, Mary E. Lassaline, and Martin J. Dennis. 2000. "Causal Status as a Determinant of Feature Centrality." *Cognitive Psychology*, 41(4): 361–416. 1

Andre, Peter, Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart. 2022. "Subjective Models of the Macroeconomy: Evidence From Experts and Representative Samples." *The Review of Economic Studies*, 89(6): 2958–2991. 1, 7, 8

Andre, Peter, Ingar Haaland, Christopher Roth, and Johannes Wohlfart. 2024. "Narratives about the Macroeconomy." *Working Paper*. 7

Andre, Peter, Philipp Schirmer, and Johannes Wohlfart. 2024. "Mental Models of the Stock Market." *Working Paper*. 7

Angeletos, George-Marios, and Chen Lian. 2018. "Forward Guidance without Common Knowledge." *American Economic Review*, 108(9): 2477–2512. 1, 5, 6, 7

Angeletos, George-Marios, and Chen Lian. 2023*a*. "Dampening General Equilibrium: Incomplete Information and Bounded Rationality." In *Handbook of Economic Expectations*. 613–645. 6, 18

Angeletos, George-Marios, and Chen Lian. 2023*b*. "Determinacy without the Taylor Principle." *Journal of Political Economy*, 131(8): 2125–2164. 30

Angeletos, George-Marios, and Zhen Huo. 2021. "Myopia and Anchoring." *American Economic Review*, 111(4): 1166–1200. 7

Arad, Ayala, and Ariel Rubinstein. 2012. "The 11–20 Money Request Game: A Level-k Reasoning Study." *American Economic Review*, 102(7): 3561–3573. 21

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2024. "The Intertemporal Keynesian Cross." *Journal of Political Economy*, , (Accepted). 36

Barberis, Nicholas, Andrei Shleifer, and Robert W. Vishny. 1998. "A Model of Investor Sentiment." *Journal of Financial Economics*, 49(3): 307–343. 7

Bastianello, Francesca, and Paul Fontanier. 2024. "Expectations and Learning from Prices." *Review of Economic Studies*, , (Forthcoming). 7

Bauer, Michael D., Carolin E. Pflueger, and Adi Sunderam. 2024. "Changing Perceptions and Post-Pandemic Monetary Policy." *Proceedings - Economic Policy Symposium - Jackson Hole*. 1, 2, 4, 7, 26, 32

Ben Zeev, Nadav, Valerie A. Ramey, and Sarah Zubairy. 2023. "Do Government Spending Multipliers Depend on the Sign of the Shock?" *AEA Papers and Proceedings*, 113: 382–387. 53

Bianchi-Vimercati, Riccardo, Martin Eichenbaum, and Joao Guerreiro. 2024. "Fiscal Stimulus with Imperfect Expectations: Spending vs. Tax Policy." *Journal of Economic Theory*, 217: 105814. 7

Blanchard, Olivier, and Roberto Perotti. 2002. "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output." *The Quarterly Journal of Economics*, 117(4): 1329–1368. 53

Blanchard, Olivier J., and Jordi Galí. 2010. "The Macroeconomic Effects of Oil Price Shocks: Why Are the 2000s so Different from the 1970s?" In *International Dimensions of Monetary Policy*. , ed. Jordi Galí and Mark Gertler, 373–421. Chicago: University of Chicago Press / (National Bureau of Economic Research Conference Report). 53

Boivin, Jean, Michael T. Kiley, and Frederic S. Mishkin. 2010. "How Has the Monetary Transmission Mechanism Evolved Over Time?" In *Handbook of Monetary Economics*. Vol. 3, 369–422. Elsevier Ltd. 53

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *American Economic Review*, 110(9): 2748–2782. 7

Breza, Emily, Arun Chandrasekhar, and Alireza Tahbaz-Salehi. 2018. "Seeing the Forest for the Trees? An Investigation of Network Knowledge." *NBER Working Paper*, w24359. 19

Camerer, Colin, and Dan Lovallo. 1999. "Overconfidence and Excess Entry: An Experimental Approach." *American Economic Review*, 89(1): 306–318. 7

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *The Quarterly Journal of Economics*, 119(3): 861–898. 7

Cieslak, Anna. 2018. "Short-Rate Expectations and Unexpected Returns in Treasury Bonds." *The Review of Financial Studies*, 31(9): 3265–3306. 2, 5, 7, 31, 32

Cloyne, James, Ezgi Kurt, and Paolo Surico. 2023. "Who Gains from Corporate Tax Cuts?" *NBER Working Paper*, w31278. 53

Cloyne, James, Joseba Martinez, Haroon Mumtaz, and Paolo Surico. 2023. "Do Tax Increases Tame Inflation?" *AEA Papers and Proceedings*, 113: 377–381. 53

Cloyne, James, Joseba Martinez, Haroon Mumtaz, and Paolo Surico. 2024. "Taxes, Innovation and Productivity." *Working Paper*. 53

Cochrane, John H., and Monika Piazzesi. 2002. "The Fed and Interest Rates—A High-Frequency Identification." *American Economic Review*, 92(2): 90–95. 2, 4, 29, 32

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, and Maarten Van Rooij. 2023*a*. "How Does Consumption Respond to News about Inflation? Field Evidence from a Randomized Control Trial." *American Economic Journal: Macroeconomics*, 15(3): 109–152. 21

Coibion, Olivier, Yuriy Gorodnichenko, Edward S. Knotek, and Raphael Schoenle. 2023*b*. "Average Inflation Targeting and Household Expectations." *Journal of Political Economy Macroeconomics*, 1(2): 403–446. 1

Coibion, Olivier, Yuriy Gorodnichenko, Saten Kumar, and Jane Ryngaert. 2021. "Do You Know That I Know That You Know. . . ? Higher-Order Beliefs in Survey Data." *The Quarterly Journal of Economics*, 136(3): 1387–1446. 21

Corsetti, Giancarlo, André Meier, and Gernot J. Müller. 2012. "Fiscal Stimulus with Spending Reversals." *Review of Economics and Statistics*, 94(4): 878–895. 53

D'Acunto, Francesco, Daniel Hoang, and Michael Weber. 2022. "Managing Households' Expectations with Unconventional Policies." *The Review of Financial Studies*, 35(4): 1597–1642. 1

Dal Bó, Ernesto, Pedro Dal Bó, and Erik Eyster. 2018. "The Demand for Bad Policy When Voters Underappreciate Equilibrium Effects." *The Review of Economic Studies*, 85(2): 964–998. 1, 21

da Silveira, Rava Azeredo, Yeji Sung, and Michael Woodford. 2024. "Optimally Imprecise Memory and Biased Forecasts." *American Economic Review*, , (Forthcoming). 7

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael Walker. 2022. "General Equilibrium Effects of Cash Transfers: Experimental Evidence From Kenya." *Econometrica*, 90(6): 2603–2643. 53

Eusepi, Stefano, and Bruce Preston. 2018. "The Science of Monetary Policy: An Imperfect Knowledge Perspective." *Journal of Economic Literature*, 56(1): 3–59. 6

Evans, George W., and Seppo Honkapohja. 2001. *Learning and Expectations in Macroeconomics. Frontiers of Economic Research*, Princeton:Princeton University Press. 6

Farhi, Emmanuel, and Iván Werning. 2019. "Monetary Policy, Bounded Rationality, and Incomplete Markets." *American Economic Review*, 109(11): 3887–3928. 1, 5, 6, 7, 30, 36

Gabaix, Xavier. 2020. "A Behavioral New Keynesian Model." *American Economic Review*, 110(8): 2271–2327. 1, 5, 7, 30

Gabaix, Xavier, and David I. Laibson. 2001. "The 6D Bias and the Equity-Premium Puzzle." *NBER Macroeconomics Annual*, 16: 257–312. 6

Galí, Jordi. 2015. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications - Second Edition.* Princeton University Press. 2, 8

Galí, Jordi, J. David López-Salido, and Javier Vallés. 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association*, 5(1): 227–270. 53

García-Schmidt, Mariana, and Michael Woodford. 2019. "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis." *American Economic Review*, 109(1): 86–120. 1, 6, 36

Georganas, Sotiris, Paul J. Healy, and Roberto A. Weber. 2015. "On the Persistence of Strategic Sophistication." *Journal of Economic Theory*, 159: 369–400. 21

Greenwood, Robin, and Samuel G. Hanson. 2015. "Waves in Ship Prices and Investment." *The Quarterly Journal of Economics*, 130(1): 55–109. 2, 6, 7

Gürkaynak, Refet S., Brian Sack, and Eric Swanson. 2005. "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models." *American Economic Review*, 95(1): 425–436. 2, 4, 29, 32

Hanson, Samuel G., and Jeremy C. Stein. 2015. "Monetary Policy and Long-Term Real Rates." *Journal of Financial Economics*, 115(3): 429–448. 2, 4, 7, 29, 32

Hanson, Samuel G., David O. Lucca, and Jonathan H. Wright. 2021. "Rate-Amplifying Demand and the Excess Sensitivity of Long-Term Rates." *Quarterly Journal of Economics*, 136(3): 1719–1781. 29

Iovino, Luigi, and Dmitriy Sergeyev. 2023. "Central Bank Balance Sheet Policies Without Rational Expectations." *Review of Economic Studies*, 90(6): 3119–3152. 7

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton. 2014. "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks." *The Journal of Finance*, 69(3): 1197–1233. 2, 5, 7, 31, 32

Känzig, Diego R. 2021. "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements." *American Economic Review*, 111(4): 1092–1125. 53

Kendall, Chad, and Ryan Oprea. 2024. "On the Complexity of Forming Mental Models." *Quantitative Economics*, 15(1): 175–211. 1

Kilian, Lutz, and Robert J. Vigfusson. 2011. "Are the Responses of the U.S. Economy Asymmetric in Energy Price Increases and Decreases?: Are Responses of the U.S. Economy Asymmetric?" *Quantitative Economics*, 2(3): 419–453. 53

Lucas, Jr., Robert E. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4(2): 103–124. 6

Lusardi, Annamaria, and Olivia S. Mitchell. 2011. "Financial Literacy and Retirement Planning in the United States." *Journal of Pension Economics and Finance*, 10(4): 509–525. 22

Maćkowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99(3): 769–803. 6

Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *The Quarterly Journal of Economics*, 117(4): 1295–1328. 6

Mendes, Arthur, Wataru Miyamoto, Thuy Lan Nguyen, Steven Pennings, and Leo Feler. 2024. "The Macroeconomic Effects of Cash Transfers: Evidence from Brazil." *Working Paper*. 53

Mertens, Karel, and Morten O. Ravn. 2013. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review*, 103(4): 1212–1247. 53

Miao, Jianjun, Jieran Wu, and Eric R. Young. 2022. "Multivariate Rational Inattention." *Econometrica*, 90(2): 907–945. 6

Miranda-Agrippino, Silvia, and Giovanni Ricco. 2021. "The Transmission of Monetary Policy Shocks." *American Economic Journal: Macroeconomics*, 13(3): 74–107. 53

Molavi, Pooya. 2019. "Macroeconomics with Learning and Misspecification: A General Theory and Applications." *Working Paper*. 6

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *The American Economic Review*, 85(5): 1313–1326. 7, 21

Nimark, Kristoffer. 2008. "Dynamic Pricing and Imperfect Common Knowledge." *Journal of Monetary Economics*, 55(2): 365–382. 6

Olivei, Giovanni, and Silvana Tenreyro. 2007. "The Timing of Monetary Policy Shocks." *American Economic Review*, 97(3): 636–663. 53

Pennings, Steven. 2021. "Cross-Region Transfer Multipliers in a Monetary Union: Evidence from Social Security and Stimulus Payments." *American Economic Review*, 111(5): 1689–1719. 53

Ramey, Valerie A. 2011. "Identifying Government Spending Shocks: It's All in the Timing." *The Quarterly Journal of Economics*, 126(1): 1–50. 53

Ramey, Valerie A. 2013. "Government Spending and Private Activity." In *Fiscal Policy after the Financial Crisis*. , ed. Alberto Alesina and Francesco Giavazzi, 19–55. University of Chicago Press. 53

Ramey, Valerie A. 2016. "Macroeconomic Shocks and Their Propagation." In *Handbook of Macroeconomics*. Vol. 2, 71–162. Elsevier B.V. 53

Rottman, Benjamin Margolin, and Reid Hastie. 2014. "Reasoning about Causal Relationships: Inferences on Causal Networks." *Psychological Bulletin*, 140(1): 109–139. 1

Rozenblit, Leonid, and Frank Keil. 2002. "The Misunderstood Limits of Folk Science: An Illusion of Explanatory Depth." *Cognitive Science*, 26(5): 521–562. 1

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3): 665–690. 6

Stahl, II, Dale O., and Paul W. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." *Journal of Economic Behavior & Organization*, 25(3): 309–327. 7

Stahl, II, Dale O., and Paul W. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, 10(1): 218–254. 7

Stantcheva, Stefanie. 2021. "Understanding Tax Policy: How Do People Reason?" *The Quarterly Journal of Economics*, 136(4): 2309–2369. 7

Stantcheva, Stefanie. 2023. "Understanding of Trade." *NBER Working Paper*. 7

Waldmann, Michael R., ed. 2017. *The Oxford Handbook of Causal Reasoning*. Vol. 1, Oxford University Press. 43

Woodford, Michael. 2003*a*. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*. , ed. Philippe Aghion, Roman Frydman, Joseph E. Stiglitz and Michael Woodford, 25–58. Princeton:Princeton University Press. 6

Woodford, Michael. 2003*b*. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press. 2, 8

Wu, Lingxuan. 2023. "Mental Macro-Finance Models: Evidence and Theory." *SSRN Electronic Journal*. 8, 13

A Details of the New Keynesian Economy

[This whole section is to be cleaned up.]

A.1 The Infinite-Horizon New Keynesian Model

Firms. There is a continuum of firms indexed by $j \in [0,1]$ in this economy subject to Calvo price rigidity. Each firm chooses labor demand N^s_{jt}, pays dividend DIV_{jt}, and sets price P^*_{jt} when it can, taking as given the aggregate inflation rate π_t, the real wage W_t, and the aggregate demand C_t. They agree on the steady state of the economy but may have heterogeneous beliefs about the economy's response to shocks.

Each firm produces a differentiated good, using the same constant-returns production technology using labor hours $Y_{jt} = N^d_{jt}$, which together forms a bundle with constant elasticity ε of substitution (CES) that the households consume. At the steady state, they each charge a markup $\mu \equiv \frac{\varepsilon}{\varepsilon-1}$. The log-linearized real dividend and aggregate labor demand are

$$div_t = c_t - \frac{1}{\mu - 1} w_t \tag{A1}$$

$$n^d_t = c_t \tag{A2}$$

Each firm resets its price with independent probability $1-\theta$ in any period and fulfills its demand period by period. Firms that can reset prices choose

$$p^*_{jt} = p_{t-1} + (1-\beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \, \mathbb{E}_{it} \left[\sum_{l=0}^{k} \pi_{t+l} + w_{t+k} \right]$$

We assume that the average belief of firms is the same as that of the households $\overline{\mathbb{E}}_t$. Aggregate inflation emerges from the pricing behavior of the $(1-\theta)$ share of resetting firms as $\pi_t = (1-\theta)\left(p^*_{jt} - p_{t-1}\right)$. Following the tradition, we consider a cost-push shock ϵ^π_t for inflation

$$\pi_t = \theta\kappa w_t + (1-\theta)\pi_t + \sum_{k=1}^{\infty} (\beta\theta)^k \, \overline{\mathbb{E}}_t [w_{t+k}] + (1-\theta) \sum_{k=1}^{\infty} (\beta\theta)^k \, \overline{\mathbb{E}}_t [\pi_{t+k}] + \epsilon^\pi_t \tag{A3}$$

with $\kappa \equiv \frac{(1-\theta)(1-\beta\theta)}{\theta}$ capturing the slope of the Phillips curve. Importantly, we do not move

π_t on the right-hand side to the left. We intentionally preserve the dependence of π_t on itself, which encapsulates the within-period complementarity in individual price setting as each firm takes the aggregate inflation as given.

Households. There is a continuum of infinitely lived households indexed by $h \in [0,1]$. Each household chooses consumption C_{ht} and labor supply N^s_{ht}, taking as given the gross nominal interest rates R_{t-1}, the inflation rate $\pi_t \equiv P_t/P_{t-1} - 1$, the real wage W_t, and the real dividend DIV_t. They agree on the steady state of the economy but may have heterogeneous beliefs about the economy's response to shocks.

They each maximize

$$\max_{\{C_{ht},N^s_{ht}\}_{t\geq 0}} \mathbb{E}_{h,t=0} \sum_{t=0}^{\infty} \beta^t \left(\frac{C_{ht}^{1-\sigma} - 1}{1-\sigma} - \frac{\left(N^s_{ht}\right)^{1+\nu}}{1+\nu} \right)$$

subject to the budget constraint

$$C_{ht} + A_{ht} = \frac{R_{t-1}}{1+\pi_t} A_{h,t-1} + W_t N^s_{ht} + DIV_t$$

where C_{ht} is a CES bundle of goods in the economy, A_{ht} is the period-t saving.

Around the steady state with zero savings, the familiar log-linearized aggregate consumption and labor supply functions are (Appendix A contains all derivations)

$$\begin{aligned} c_t = & -\sigma^{-1}\beta i_t - \sigma^{-1}\beta \sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}\left[i_{t+k}\right] + \sigma^{-1} \sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}\left[\pi_{t+k}\right] + (1-\beta)\left[\frac{(\mu-1)\,\nu}{\sigma+\mu\nu} div_t + \frac{(1+\nu)}{\sigma+\mu\nu} w_t\right] \\ & + (1-\beta)\sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}\Big[\frac{(\mu-1)\,\nu}{\sigma+\mu\nu} div_{t+k} + \frac{(1+\nu)}{\sigma+\mu\nu} w_{t+k}\Big] \end{aligned} \tag{A4}$$

$$\begin{aligned} n^s_t = & \,\nu^{-1}\beta i_t + \nu^{-1}\beta \sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}\left[i_{t+k}\right] - \nu^{-1} \sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}\left[\pi_{t+k}\right] - (1-\beta)\,\frac{(\mu-1)\,\sigma}{\sigma+\mu\nu} div_t + \nu^{-1}\left(1 - \sigma\frac{(1-\beta)\,(1+\nu)}{\sigma+\mu\nu}\right) w_t \\ & - (1-\beta)\sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}\Big[\frac{(\mu-1)\,\sigma}{\sigma+\mu\nu} div_{t+k} + \frac{(1+\nu)\,\sigma\nu^{-1}}{\sigma+\mu\nu} w_{t+k}\Big] \end{aligned} \tag{A5}$$

where $\overline{\mathbb{E}}_t[\cdot]$ denotes the average expectations of households, the lower-case variables denote the log deviation from the corresponding steady-state values, and $\mu = \frac{C}{WN^s}$ denotes the ratio of consumption to labor income at the steady state (which equals firms' markups as introduced next).

Monetary policy. The central bank a Taylor rule with a monetary policy shock ϵ_t^i,

$$i_t = \phi \pi_t + \epsilon_t^i \tag{A6}$$

Competitive labor market. Last, to close the model, the wage arises by equilibrating labor supply and demand

$$n_t^s = n_t^d \tag{A7}$$

A.2 Derivations

Proof of Proposition 3.

Proof of Proposition 4.

B Details of the RBC Economy

C Additional Figures and Tables

C.1 Tables for the Forecast Part of the Survey

Table C1: Variables elicited in forecast part of our survey

		Group 1		Group 2		Group 3	
Variable	Abbrev.	Oil ↑	MP ↑	G ↑	PIT ↑	CIT ↑	TP ↑
Firms-related	**bus**						
Nominal marginal cost	mc	✓	✓	✓	✓	✓	✓
Demand	Y	✓	✓	✓	✓	✓	✓
Interest rate	i	✓	✓	✓	✓	✓	✓
Corporate income tax rate	CIT	✓	✓	✓	✓	✓	✓
Prices	p	✓	✓	✓	✓	✓	✓
Intermediate inputs	x	✓	✓	✓	✓	✓	✓
Investment	I	✓	✓	✓	✓	✓	✓
Total hours	N	✓	✓	✓	✓	✓	✓
Unemployment rate	u	✓	✓	✓	✓	✓	✓
Post-tax profits	d	✓	✓	✓	✓	✓	✓
Households-related	**hh**						
Interest rate	i	✓	✓	✓	✓	✓	✓
Prices	p	✓	✓	✓	✓	✓	✓
Hours	N	✓	✓	✓	✓	✓	✓
Personal income tax rate	PIT	✓	✓	✓	✓	✓	✓
Pre-tax nominal wage	W	✓	✓	✓	✓	✓	✓
Durable consumption	D	✓	✓	✓	✓	✓	✓
Non-durable consumption	ND	✓	✓	✓	✓	✓	✓
Central-bank-related	**fed**						
Unemployment rate	u	✓	✓	✓	✓	✓	✓
Inflation	p	✓	✓	✓	✓	✓	✓
Interest rate	i	✓	×	✓	✓	✓	✓
Government-related	**gov**						
Borrowing/repayment	B	✓	✓	✓	✓	✓	✓
Tax revenue	TR	✓	✓	✓	✓	✓	✓

Table C2: Literature review of directional impulse responses

		Oil price ↑	Monetary policy (MP) ↑	Transfer payment (TP)‡ ↑
Output	Y	Down (Känzig)	Down (Ramey)	Up (Pennings)
Interest rate	i	Up (Känzig)	Up (Ramey, as shock itself)	
Price	p	Up (Känzig)	Down (MAR), insignificant or up (Ramey, classic price puzzle)	
Unemployment	u	Up (Känzig)	Up (Ramey)	Down (MMNPF)
Labor hours	N	Down (BG)	Down (MAR)	
Nonresidential investment	I	Down (Känzig)	Down (BKM)	
Durable consump.	D	Down (Känzig)	Down (BKM)	Up (EHMNW)
Nondurables & services	ND	Down (Känzig)	Down (BKM)	Up (EHMNW)
Nominal wage	W	Up (BG)	Down (OT, as real wage and price both down), insignificantly down (MAR)	Up (EHMNW)
Business post-tax profits	div			Up (EHMNW)
References		Känzig (2021, figs 8-10/A.7) Blanchard and Galí (2010, fig 7.6.A)	Ramey (2016, figs 2-3) Miranda-Agrippino and Ricco (2021, fig 7) Boivin, Kiley and Mishkin (2010, fig 4, post-84) Olivei and Tenreyro (2007, figs 10-14)	Pennings (2021, tab 1) Mendes et al. (2024, tab 5) Egger et al. (2022, tabs 1, 3)

		Government spending (G) ↑	Personal income tax (PIT) ↑	Corporate income tax (CIT) ↑
Output	Y	Up (Ramey16)	Down (MR)	Down (MR)
Interest rate	i	Insignificant (Ramey11)	Insignificant (MR)	Insignificant (MR)
Price	p	Up (Ramey16 unreported result)	Down (CMMS23), insignificantly down (MR)	Insignificantly up (CMMS23, MR)
Unemployment	u	Down (Ramey13*)	Up (MR)	Up (CKS), insignificantly up (MR)
Labor hours	N	Up (Ramey11*)	Down (CMMS24, MR)	Insignificantly down (CMMS24), insignificant (MR)
Nonresidential investment	I	Down (Ramey11)	Down (MR)	Down (MR)
Durable consump.	D	Down (Ramey11**)	Down (MR)	Down (CMMS24†), insignificantly up (MR)
Nondurables & services	ND	Down (Ramey11**)	Down (CMMS24†), insignificantly down (MR)	Down (CMMS24†), insignificantly up (MR)
Nominal wage	W	Up (Ramey11)	Down (CMMS24, as real wage and price both down)	Down (CMMS24, as real wage down and price insignificant)
Business post-tax profits	div			Down (if measured by dividend, CKS)
Personal income tax rate	PIT	Up (Ramey11)	Up (MR, as shock itself)	Insignificantly down (MR)
Corporate income tax rate	CIT	Up (Ramey16***)	Insignificantly down (MR)	Up (MR, as shock itself)
Tax revenue	TR		Up (CMMS24‡, MR)	Up (CMMS24††), insignificant (MR)
Government spending	G	Up (Ramey16, as shock itself)	Insignificant (MR)	Insignificant (MR)
Government debt	B	Up (CMM)	Down (CMMS24, MR)	Insignificant (CMMS24, MR)
References		Ramey (2016, fig 5) Ramey (2011, fig X) Ramey (2013, figs 1.11-17) Corsetti, Meier and Müller (2012, fig 1)	Mertens and Ravn (2013, figs 2-4/9/10) Cloyne et al. (2023, fig 1) Cloyne et al. (2024, figs 1-2/B.1/H.8) Cloyne, Kurt and Surico (2023, figs 2/3B)	

Notes: This table lists directional impulse responses at about 1-year horizon across variables to shocks. Shocks considered are oil shocks, contractionary monetary policy (MP) shocks, transfer payment (TP) shocks, government spending (G) shocks, and positive shocks in personal income tax (PIT) rate and corporate income tax (CIT) rate. As we are only interested in the directions rather than the magnitudes of responses, we only impose a weak assumption that each variable responds to positive and negative shocks with opposite signs. This is weaker than assuming that the multipliers are the same for positive and negative shocks. Nonetheless, it is worth noting that while some earlier papers advocate for asymmetric multipliers, more recent papers argue that the evidence is weak (e.g., Kilian and Vigfusson (2011) for oil shocks and Ben Zeev, Ramey and Zubairy (2023) for government spending shocks). The abbreviation in parentheses indicates the main reference, usually the most recent or most cited paper. All references are listed at the bottom of each column.

‡ For the transfer payment shock, Pennings (2021) and Mendes et al. (2024) provide cross-sectional estimates instead of aggregate ones, in the US and Brazil respectively. Egger et al. (2022) study transfers in Kenya, but these transfers are funded from outside the economy. We drop this shock in a robustness version in Table C3.

* Ramey (2013) shows that the unemployment rate falls in response to government spending shocks. It is mainly driven by government employment, with the response of private employment either insignificant or negative in different specifications. We drop this variable in a robustness version in Table C3, since we elicit participants' opinions about private businesses.

** Ramey (2011, 2016) discusses extensively potential issues with previous works (e.g., Blanchard and Perotti, 2002; Galí, López-Salido and Vallés, 2007) that find a positive consumption response to government spending shocks. We drop this variable in a robustness version in Table C3.

*** Ramey (2016, fig 5) suggests that the average tax rate goes up in response to government spending shocks, calculated as federal current receipts divided by nominal GDP.

† Cloyne et al. (2024, fig 2) show positive consumption responses to decreases in PIT and CIT, but do not split into durables vs. nondurables.

†† Cloyne et al. (2024, fig H.8) show negative primary surplus response (tax revenue minus government spending) to decreases in PIT and CIT.

Table C3: Correct directional responses with robustness versions

	Oil ↑			MP ↑			G ↑			PIT ↑			CIT ↑			TP ↑		
Version	Va	Vb	Vc	Va	Vb	Vc	Va	Vb	Vc	Va	Vb	Vc	Va	Vb	Vc	Va	Vb	Vc
Y	↓	↓	↓	↓	↓	↓	↑	↑	↑	↓	↓	↓	↓	↓	↓	↑	↑	↑
i	↑	↑	↑	↑	↑	↑			↑			↓			↓			↑
p	↑	↑	↑	↓		↓	↑		↑	↓		↓			↓			↑
u	↑	↑	↑	↑	↑	↑	↓		↓	↑	↑	↑	↑		↑	↓	↓	↓
N	↓	↓	↓	↓	↓	↓	↑	↑	↑	↓	↓	↓	↓		↓	↑	↑	↑
I	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓			
D	↓	↓	↓	↓	↓	↓	↓		↓	↓	↓	↓	↓		↓	↑		↑
ND	↓	↓	↓	↓	↓	↓	↓		↓	↓		↓	↓		↓	↑		↑
W	↑	↑	↑	↓	↓	↓	↑	↑	↑	↓	↓	↓	↓	↓	↓	↑		↑
div			↓			↓			↑			↓	↓	↓	↓	↑		↑
mc			↑			↓			↑			↓			↓			↑
PIT							↑	↑	↑	↑	↑	↑						↑
CIT							↑		↑				↑	↑	↑			↑
TR									↑	↑	↑	↑	↑		↑			↑
B							↑	↑	↑	↓	↓	↓	↓		↓			↑

Notes: Three versions are constructed as follows. Version a is the baseline specification in Table 1.
Version a is most closely based on most up-to-date empirical estimates with clear directions.
Version b drops from Version a variables for which estimates are noisy or controversy exists. A few noteworthy choices are price response to monetary shocks (classic price puzzle), unemployment and consumption responses to government spending shocks (as discussed in the notes of Table C2).
Version c makes additional predictions based on theoretical predictions, relative to Version a.

Table C4: Model-implied variable depth with robustness versions

	Oil			MP			G			PIT			CIT			TP		
Model	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
Y	3	3	3	2	2	2	1	1	1	2	2	2	3	3	3	2	2	2
i	2	2	2	1	1	1	5	3	3	5	3	4	6	4	3	5	4	3
p	1	1	1	4	3	4	4	2	2	4	3	3	5	4	2	4	3	2
u	3	3	3	2	2	2	2	2	2	2	2	2	3	3	2	2	3	2
N	3	3	3	2	2	2	2	2	2	2	2	2	3	3	2	2	3	2
I			3			2			2			3			2			2
D	3	3	3	2	2	2	2	2	2	2	2	2	3	3	3	2	2	2
ND	3	3	3	2	2	2	2	2	2	2	2	2	3	3	3	2	2	2
W			2			3			3			3			3			3
div	4	4	3	3	3	3	2	2	2	3	3	3	2	2	2	2	2	2
mc	1	1	1	3	3	3	3	2	2	3	3	3	4	4	2	3	3	2
PIT							1	1	1	1	1	1				1	1	1
CIT							1	1	1				1	1	1	1	1	1
B							1	1	1	1	1	1	1	1	1	1	1	1

Notes: Three sets of depths are constructed as follows, with the oil price shock interpreted as a cost-push shock. Model 2 is the baseline specification in Table 1.
Model 1 is the textbook New Keynesian model in the main text of this paper.
Model 2 extends Model 1 to feature decreasing-returns production (so that the marginal cost is increasing in quantity) and a Taylor rule of monetary policy that depends on both inflation and unemployment.
Model 3 extends Model 1 with capital investment by firms, price and wage rigidity (via a labor union instead of a competitive labor market).

C.2 Additional Results for Fact 1