A BEHAVIORAL NEW KEYNESIAN MODEL

Xavier Gabaix

A Behavioral New Keynesian Model
Xavier Gabaix
NBER Working Paper No. 22954
December 2016, Revised June 2019
JEL No. D03,E03,E5,E6,G02

ABSTRACT

This paper presents a framework for analyzing how bounded rationality affects monetary and fiscal policy. The model is a tractable and parsimonious enrichment of the widely-used New Keynesian model – with one main new "cognitive discounting" parameter, which quantifies how poorly agents understand future economic disturbances. That myopia parameter, in turn, affects the power of monetary and fiscal policy in a microfounded general equilibrium. A number of consequences emerge. (i) Fiscal stimulus or "helicopter drops of money" are powerful and, indeed, pull the economy out of the zero lower bound. More generally, the model allows for the joint analysis of optimal monetary and fiscal policy. (ii) The Taylor principle is strongly modified: even with passive monetary policy, equilibrium is determinate, whereas the traditional rational model yields multiple equilibria, which reduces its predictive power, and generates indeterminate economies at the zero lower bound (ZLB). (iii) The ZLB is much less costly than in the traditional model. (iv) The model brings a natural solution to the "forward guidance puzzle": the fact that in the rational model, shocks to very distant rates have a very powerful impact on today's consumption and inflation; because agents are partially myopic, this effect is muted. (v) Optimal policy changes qualitatively: the optimal commitment policy with rational agents demands "nominal GDP targeting"; this is not the case with behavioral firms, as the benefits of commitment are less strong with myopic firms. (vi) The model is "neo-Fisherian" in the long run, but Keynesian in the short run: a permanent rise in the interest rate decreases inflation in the short run but increases it in the long run. The non-standard behavioral features of the model seem warranted by extant empirical evidence.

Xavier Gabaix
Department of Economics
Harvard University
Littauer Center
1805 Cambridge St
Ḉambridge, MA 02138
and NBER
xgabaix@fas.harvard.edu

1   Introduction

This paper proposes a way to analyze monetary and fiscal policy when agents are not fully rational. To do so, it enriches the basic model of monetary policy, the New Keynesian (NK) model, by incorporating behavioral factors. In the baseline NK model the agent is fully rational (though prices are sticky). Here, in contrast, the agent is partially myopic to unusual events and does not anticipate the future perfectly. The formulation takes the form of a parsimonious generalization of the traditional model that allows for the analysis of monetary and fiscal policy. This has a number of strong consequences for aggregate outcomes.

1. Fiscal policy is much more powerful than in the traditional model.[1] In the traditional model, rational agents are Ricardian and do not react to tax cuts. In the present behavioral model, agents are partly myopic, and consume more when they receive tax cuts or "helicopter drops of money" from the central bank. As a result, we can study the interaction between monetary and fiscal policy.

2. The Taylor principle is strongly modified. Equilibrium selection issues vanish in many cases: for instance, even with a constant nominal interest rate there is just one (bounded) equilibrium.

3. Relatedly, the model can explain the stability of economies stuck at the zero lower bound (ZLB), something that is difficult to achieve in traditional models.

4. The ZLB is much less costly.

5. Forward guidance is much less powerful than in the traditional model, offering a natural behavioral resolution of the "forward guidance puzzle".

6. Optimal policy changes qualitatively: for instance, the optimal commitment policy with rational firms demands "price level targeting". This is not the case with behavioral firms.

7. A number of neo-Fisherian paradoxes are resolved. A permanent rise in the nominal interest rate causes inflation to fall in the short run (a Keynesian effect) and rise in the long run (so that long-run Fisher neutrality holds with respect to inflation).

In addition, I will argue that there is reasonable empirical evidence for the main non-standard features of the model. Let me expand on the above points.

Behavioral mechanism: Cognitive discounting. The main non-standard feature of the model is a form of *cognitive discounting.* The phenomenon I want to analyze is that the world is not fully understood by agents, especially events that are far into the future. To capture this mathematically, I make the following key assumption: agents, as they simulate the future, shrink their simulations toward a simple benchmark — namely, the steady state of the economy.[2] This follows from a microfoundation in which agents receive noisy signals about the economy. As a result, an innovation happening in k periods has a direct impact on agents' expectations that is shrunk by a factor $\bar{m}^k$ relative to the rational response, where $\bar{m} \in [0, 1]$ is a parameter capturing cognitive discounting. Hence, innovations that are deep in

---

[1] By "fiscal policy" I mean government transfers, i.e. changes in (lump-sum) taxes. In the traditional Ricardian model, they have no effect (Barro (1974)). This is in contrast to government consumption, which does have an effect even in the traditional model.

[2] The paper also explores a variant in which the simulations are biased toward the contemporaneous values. The conclusions are similar: future impacts are heavily discounted.

the future get heavily discounted relatively to a rational benchmark, which corresponds to the case $\bar{m} = 1$. The agent is globally patient with respect to steady-state variables, but is myopic with respect to deviations around the steady state, especially if they are in the distant future.

This one premise leads to a behavioral version of the agents' Euler equation and of the firms' price-setting policy. In general equilibrium, this yields the following tractable behavioral amendment of the canonical two-equation New Keynesian model (which is fully rational and is distilled in Woodford (2003b) and Galí (2015)), in which the future output gap and inflation enter in a discounted way:

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] - \sigma\left(i_t - \mathbb{E}_t\pi_{t+1} - r_t^n\right) \text{ (IS curve)}, \tag{1}$$

$$\pi_t = \beta M^f \mathbb{E}_t\left[\pi_{t+1}\right] + \kappa x_t \text{ (Phillips curve)}, \tag{2}$$

where x_t is the output gap (the deviation of GDP from its efficient level, so that positive x_t corresponds to a boom, negative x_t to a recession), π_t is inflation, i_t is the nominal interest rate, r_t^n is the natural real interest rate, σ is the sensitivity of the output gap to the interest rate, κ is the sensitivity of inflation to the output gap, and β is the pure rate of time preference. The equilibrium behavioral parameters $M, M^f \in [0, 1]$ are the aggregate-level attention parameters of consumers and firms, respectively, and increase with the micro cognitive discounting parameter $\bar{m}$. The traditional, rational model corresponds to $M = M^f = 1$, but the behavioral model generates M and M^f strictly less than 1. This has a large number of consequences for macroeconomic dynamics.

Fiscal policy and helicopter drops of money. In the traditional NK model, agents are fully rational. So Ricardian equivalence holds, and fiscal policy (meaning lump-sum tax changes, as opposed to government expenditure) has no impact. Here, in contrast, the agent is not Ricardian because he fails to perfectly anticipate future taxes. As a result, tax cuts and transfers are unusually stimulative (technically, they increase the natural rate r_t^n), particularly if they happen in the present. As the agent is partially myopic, taxes are best enacted in the present.

At the ZLB, only forward guidance (or, in more general models, quantitative easing) is available, and in the rational model optimal policy only leads to a complicated second best. However, in this model, the central bank (and more generally the government) has a new instrument: it can restore the first best by doing "helicopter drops of money", i.e. by sending checks to people – via fiscal policy.

Zero lower bound (ZLB). Depressions due to the ZLB are unboundedly large in the rational model, probably counterfactually so (see for example Werning (2012)). This is because agents unflinchingly respect their Euler equations. In contrast, depressions are moderate and bounded in this behavioral model – closer to reality and common sense. This is because aggregate discounting ($M, M^f < 1$) induces exponential attenuation of the impact of future output gaps on current macroeconomic conditions.

The Taylor principle reconsidered and equilibrium determinacy. When monetary policy is passive (e.g. via a constant interest rate rule, or when it violates the Taylor principle that monetary policy should strongly lean against economic conditions), the traditional model has a continuum of (bounded) equilibria, so that the response to a simple question like "What happens when interest rates are kept constant?" is ill-defined: it is mired in the morass of equilibrium selection. In contrast, in this behavioral model there is just one (bounded) equilibrium: things are clean and theoretically definite. A partial intuition is that the discounting parameters M and M^f make the present less sensitive to the future and hence stabilize the economy (in the technical sense of making its equilibrium "Blanchard and Kahn (1980) determinate").

Economic stability. Determinacy is not just a purely theoretical question. In the rational model, if

3

the economy is stuck at the ZLB forever the Taylor principle is violated (as the nominal interest rate is stuck at 0%). The equilibrium is therefore indeterminate: we could expect the economy to jump randomly from one period to the next (we shall see that a similar phenomenon happens if the ZLB lasts for a large but finite duration). However, we do not see that in Japan since the late 1980s or in Western economies in the aftermath of the 2008 crisis (Cochrane (2017)). This can be explained with this behavioral model if agents are myopic enough and if firms rely enough on "inflation guidance" by the central bank.

Forward guidance. With rational agents, "forward guidance" by the central bank is predicted to work very powerfully, most likely too much so, as emphasized by Del Negro et al. (2015) and McKay et al. (2016). The reason is again that the traditional consumer rigidly respects his Euler equation and expects other agents to do the same, so that a movement of the interest rate far in the future has a strong impact today. However, in the behavioral model this impact is muted by the agent's myopia $(M, M^f < 1)$, which makes forward guidance less powerful.

Optimal policy changes qualitatively. With rational firms, the optimal commitment policy entails "price level targeting" (which gives, when GDP is trend-stationary, "nominal GDP targeting"): after a cost-push shock, monetary policy should partially let inflation rise, but then create deflation, so that eventually the price level and nominal GDP come back to their pre-shock trend. This is because with rational firms, there are strong benefits from commitment to being very tough in the future (Clarida et al. (1999)). With behavioral firms, in contrast, the benefits from commitment are lower, and after the cost-push shock the central bank does not find it useful to engineer a deflation and come back to the initial price level (this is in part because forward guidance does not work as well). Hence, price level targeting and nominal GDP targeting are not desirable when firms are behavioral.

A number of neo-Fisherian paradoxes vanish. A number of authors, especially Cochrane (2017), highlight that in the rational New Keynesian model, a permanent rise in interest rates leads to an immediate rise in inflation, which is paradoxical. This is called the "neo-Fisherian" property. In the present behavioral model (augmented to allow for non-zero trend inflation), the property holds in the long run: the long-run real rate is independent of monetary policy (Fisher neutrality holds). However, in the short run, raising rates does lower inflation and output, as in the Keynesian model.

Links with the literature   This is a behavioral paper, and there are of course many ways to model bounded rationality. This includes limited information updating (Gabaix and Laibson (2002), Mankiw and Reis (2002)), related differential salience (Bordalo et al. (2013)), and noisy signals (Sims (2003), Maćkowiak and Wiederholt (2015), Caplin et al. (2017), Hébert and Woodford (2018)).[3] The modeling style and microfoundations of the present paper use the "sparsity" approach developed progressively since Gabaix (2014), because it seems to capture a good deal of the psychology of attention (Gabaix (2019)) while remaining highly tractable because it uses deterministic models (unlike models with noisy signals) and continuous parameters. As a result, it applies to microeconomic problems like basic consumer theory and Arrow-Debreu style general equilibrium (something as of yet not done by other modelling techniques), dynamic macroeconomics (Gabaix (2016)), and public economics (Farhi and Gabaix (2017)).

On the macro theory front, this paper relates to the literature on "macroeconomics without the rational-expectations hypothesis" reviewed in Woodford (2013). In previous literature, a popular way

---

[3]My notion of "behavioral" here is bounded rationality or cognitive myopia. I abstract from other interesting forces, like fairness (Eyster et al. (2017)) – they create an additional source of price stickiness.

to discipline beliefs under deviation from rational expectations is via learning, as reviewed in Eusepi and Preston (2018), and Evans and Honkapohja (2001).[4] This is also related to more recent work, such as that on level-k thinking by Farhi and Werning (2017), Woodford (2018), and García-Schmidt and Woodford (2019). There, agents are rational with respect to partial equilibrium effects, but don't quite understand general equilibrium effects. In both cases, the future is dampened. In the present framework agents' cognitive myopia applies to both partial and general equilibrium effect.

The paper also relates to the literature on incomplete information and higher-order beliefs that has followed Morris and Shin (1998); see Angeletos and Lian (2016) for a review. This literature has emphasized how lack of common knowledge and higher-order uncertainty can anchor expectations of endogenous outcomes, such as past inflation. Examples of such anchoring include Woodford (2003a), Morris et al. (2006), Nimark (2008), and Angeletos and La'O (2010). Building on this literature, Angeletos and Lian (2018) have explored a version of the NK model that maintains the rational hypothesis but allows for higher-order uncertainty. In a general-equilibrium setting, higher-order uncertainty anchors forward-looking expectations in a way that resembles cognitive discounting and under some assumptions can rationalize it. Angeletos and Lian (2017) provide an abstract framework that embeds rational models (with incomplete information) and behavioral models (without rational expectations).[5] A recent paper by Angeletos and Huo (2019) also shows that incomplete information offers a plausible rationalization of partly backward-looking behavior in the New Keynesian model, in addition to partially myopic features.[6] One can hope that a healthy interplay between incomplete information models and behavioral models will continue, as both are useful tools to investigate the complexities of economic life.

Rational models with financially constrained consumers can also generate something akin to $M < 1$ (but keep $M^f = 1$). McKay et al. (2016) provide a microfoundation for an approximate version of the IS curve with $M < 1$, based on heterogeneous rational agents with limited risk sharing (see also Campbell et al. (2017)), without an analysis of deficits. The analysis of Werning (2015) yields a modified Euler equation with rational heterogeneous agents, which often yields $M > 1$. Del Negro et al. (2015) and Eggertsson et al. (2019) work out models with finitely-lived agents that give an M slightly below 1. Finite lives severely limit how myopic agents can be (e.g., the models predict an M very close to 1), given that life expectancies are quite high. Galí (2017) shows that the NK model with finite lives without any assets in positive net supply (government debt, capital, or bubbles) cannot generate discounting in the IS curve.[7] Bilbiie (2018), analyzing a class of heterogeneous-agent rational models with financial constraints, finds that they can generate $M < 1$ (with some assumptions about the cyclicality of inequality), but they also generate a government spending multiplier less than 1 in the absence of counteracting channels. In contrast, the present behavioral model generates $M < 1$ as well as a failure of Ricardian equivalence that induces a government spending multiplier greater than

---

[4]Agents in these models act as econometricians who update their forecast models as new data arrive. This literature focuses greatly on the local stability of the learning dynamics, which links tightly to whether the learning equilibrium converges to the rational expectations equilibrium. Behavioral agents in the present work, on the other hand, are behavioral in the sense that their mental model never converges to the rational expectations solution.

[5]In earlier work, Angeletos and La'O (2009) and Angeletos et al. (2018) explore how dropping the common-prior assumption can help mimic the dynamics of higher-order beliefs introduced by incomplete information, which generates a departure from traditional rational expectations.

[6]The Mankiw and Reis (2002) model changes the Phillips curve, which helps with some paradoxes (Kiley (2016)), but it keeps the same IS curve as the traditional model. A synthesis of Mankiw-Reis and the present model would be useful.

[7]A related direction is via safe asset premia. Fisher (2015) derives a discounted Euler equation with a safe asset premium: but the effect is very small; for example the coefficient $1 - M$ is very close to 0 – close to the empirical "safety premium", so at most $M = 0.99$. Caballero and Farhi (2017) offer a different explanation of the forward guidance puzzle in a model with a shortage of safe assets and endogenous risk premia.

1. Hence, this behavioral model offers a resolution to the paradoxes with rational, credit-constrained agents highlighted by Bilbiie (2018).

Paper outline Section 2 presents basic model assumptions and derives its main building blocks, summarized in Proposition 2.5. Section 3 derives the positive implications of the model. Section 4 studies optimal monetary policy with behavioral agents. Section 5 analyzes fiscal policy and studies its optimal implementation. Section 6 introduces a number of model enrichments that extend the model in empirically plausible directions. These include variants of the model that allow agents to be inattentive to contemporaneous variables, which further lessens the impact of monetary policy and flattens the Phillips curve. In the baseline model, trend inflation is normalized to zero. The more empirically realistic case where inflation has a non-zero or even time-varying trend is also examined in Section 6; this also allows me to analyze neo-Fisherian paradoxes. Section 7 discusses the behavioral assumptions of the model at greater depth, and Section 8 concludes. Section 9 presents detailed microfoundations for the behavioral model. Section 10 presents an elementary 2-period model with behavioral agents. I recommend it to entrants to this literature. The rest of the appendix contains additional proofs and details.

Notations. I distinguish between $\mathbb{E}[X]$, the objective expectation of X, and $\mathbb{E}^{BR}[X]$, the expectation under the agent's *b*oundedly *r*ational (BR) model of the world.

Though the exposition is largely self-contained, this paper is in part a behavioral version of Chapters 2-5 of the Galí (2015) textbook, itself in part a summary of Woodford (2003b). My notations are typically those of Galí, except that γ is risk aversion, something that Galí denotes with σ. In concordance with the broader literature, I use σ for the ("effective") intertemporal elasticity of substitution.

Throughout the paper, I use the notation $\bar{r}$ for the steady state value of the real interest rate, or alternatively r where it is unambiguous to do so. I call the economy "determinate" (in the sense of Blanchard and Kahn (1980)) if, given initial conditions, there is only one non-explosive equilibrium path.

2 A Behavioral Model

2.1 Basic Setup and the Household's Problem

Setup: Objective reality I consider an agent with standard utility

$$U = \mathbb{E} \sum_{t=0}^{\infty} \beta^t u(c_t, N_t) \text{ with } u(c, N) = \frac{c^{1-\gamma}-1}{1-\gamma} - \frac{N^{1+\phi}}{1+\phi}, \tag{3}$$

where c_t is consumption, and N_t is labor supply (as in *N*umber of hours supplied).

There is a Dixit-Stiglitz continuum $[0,1]$ of firms producing intermediate goods. If the agent consumes c_{it} of good i, his aggregated consumption is $c_t = \left(\int_0^1 c_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$. Firm i produces output according to $Y_{it} = N_{it}e^{\zeta_t}$, where N_{it} is its labor input, and ζ_t an aggregate TFP level, which for concreteness follows an AR(1) process with mean 0. The final good is produced competitively in quantity $Y_t = \left(\int_0^1 Y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$. In equilibrium the labor market clears, so that $N_t = \int_0^1 N_{it} di$. So, if there is no pricing friction, the aggregate production of the economy is $c_t = e^{\zeta_t} N_t$. There is no capital, as in the baseline New Keynesian model.

The real wage is ω_t. The real interest rate is r_t and the agent's real income is $y_t = \omega_t N_t + y_t^f$: the sum of labor income $\omega_t N_t$ and profit income y_t^f (as in income coming from $f$irms); later we will add taxes. His real financial wealth k_t evolves as:

$$k_{t+1} = (1 + r_t)(k_t - c_t + y_t). \tag{4}$$

The agent's problem is $\max_{(c_t, N_t)_{t \geq 0}} U$ subject to (4), and the usual transversality condition ($\lim_{t \to \infty} \beta^t c_t^{-\gamma} k_t = 0$), which I will omit mentioning from now on.

Consider first the case where the economy is deterministic at the steady state ($\zeta_t \equiv 0$), so that the interest rate, income, real wage, consumption, and labor supply are at their steady-state values $\bar{r}$, $\bar{y}$, $\bar{\omega}$, $\bar{c}$, $\bar{N}$. We have a simple deterministic problem. Defining $R := 1 + \bar{r}$, we have $R = 1/\beta$. To correct monopolistic distortions, I assume that the government has put in place the usual corrective production subsidies, financed by a lump-sum tax on firms (so that profits are 0 on average). Hence, at the steady state the economy operates efficiently and $\bar{c} = \bar{N} = \bar{\omega} = \bar{y} = 1$.[8]

Let us now go back to the general case, outside of the steady state. There is a state vector $\boldsymbol{X}_t$ (comprising productivity ζ_t, as well as announced actions in monetary and fiscal policy), that will evolve in equilibrium as:

$$\boldsymbol{X}_{t+1} = \boldsymbol{G}^{\boldsymbol{X}}(\boldsymbol{X}_t, \boldsymbol{\epsilon}_{t+1}) \tag{5}$$

for some equilibrium transition function $\boldsymbol{G}^{\boldsymbol{X}}$ and mean-0 innovations $\boldsymbol{\epsilon}_{t+1}$ (which depend on the equilibrium policies of the agent and the government).

I decompose the values as deviations from the above steady state, for example:

$$r_t = \bar{r} + \hat{r}_t, \qquad y_t = \bar{y} + \hat{y}_t,$$

and those deviations are functions of the state:

$$\hat{r}_t = \hat{r}(\boldsymbol{X}_t), \qquad \hat{y}_t = \hat{y}(N_t, \boldsymbol{X}_t) := \omega(\boldsymbol{X}_t) N_t + y^f(\boldsymbol{X}_t) - \bar{y},$$

where the functions of $\boldsymbol{X}_t$ are determined in equilibrium. The law of motion for private financial wealth k_t is[9]

$$k_{t+1} = G^k(c_t, N_t, k_t, \boldsymbol{X}_t) := (1 + \bar{r} + \hat{r}(\boldsymbol{X}_t))(k_t + \bar{y} + \hat{y}(N_t, \boldsymbol{X}_t) - c_t), \tag{6}$$

so the agent's problem can be rewritten as $\max_{(c_t, N_t)_{t \geq 0}} U$ subject to (5) and (6).

I assume that $\boldsymbol{X}_t$ has mean 0, i.e. has been de-meaned. Linearizing, the law of motion becomes:

$$\boldsymbol{X}_{t+1} = \boldsymbol{\Gamma} \boldsymbol{X}_t + \boldsymbol{\epsilon}_{t+1} \tag{7}$$

for some matrix $\boldsymbol{\Gamma}$, after perhaps a renormalization of $\boldsymbol{\epsilon}_{t+1}$. Likewise, linearizing we will have $\hat{r}(\boldsymbol{X}) = \boldsymbol{b}_{\boldsymbol{X}}^r \boldsymbol{X}$, for some factor $\boldsymbol{b}_{\boldsymbol{X}}^r$.

---

[8]Indeed, when $\zeta = 0$, $\bar{\omega} = 1$, and labor supply satisfies $\bar{\omega} u_c + u_N = 0$, i.e. $\bar{N}^\phi = \bar{\omega} \bar{c}^{-\gamma}$, with the resource constraint: $\bar{c} = \bar{N}$.

[9]As there is no aggregate capital, financial wealth is $k_t = 0$ in equilibrium in the basic model without government debt. But we need to consider potential deviations from $k_t = 0$ when studying the agent's consumption problem. When later we add government debt B_t, we will have $k_t = B_t$ in equilibrium.

Setup: Reality perceived by the behavioral agent   I can now describe the behavioral agent. The main phenomenon I want to capture is that the world is not fully understood by the agent, especially events that are far into the future. To capture this mathematically, I make the following key assumption:[10]

Assumption 2.1 (Cognitive discounting of the state vector) *The agent perceives that the state vector evolves as:*

$$\boldsymbol{X}_{t+1} = \bar{m}\boldsymbol{G^X}\left(\boldsymbol{X}_t, \boldsymbol{\epsilon}_{t+1}\right), \tag{8}$$

where $\bar{m} \in [0,1]$ is a "cognitive discounting" parameter measuring attention to the future.

Then, given this perception, the agent solves $\max_{(c_t, N_t)_{t \geq 0}} U$ subject to (6) and (8).[11]

To better interpret $\bar{m}$, let us linearize (8):

$$\boldsymbol{X}_{t+1} = \bar{m}\left(\boldsymbol{\Gamma X}_t + \boldsymbol{\epsilon}_{t+1}\right). \tag{9}$$

Hence the expectation of the behavioral agent is $\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+1}\right] = \bar{m}\boldsymbol{\Gamma X}_t$ and, iterating, $\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k}\right] = \bar{m}^k\boldsymbol{\Gamma}^k\boldsymbol{X}_t$, while the rational expectation is $\mathbb{E}_t\left[\boldsymbol{X}_{t+k}\right] = \boldsymbol{\Gamma}^k\boldsymbol{X}_t$ (the rational policy always obtains from setting the attention parameters to 1).[12] Hence:

$$\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k}\right] = \bar{m}^k\mathbb{E}_t\left[\boldsymbol{X}_{t+k}\right], \tag{10}$$

where $\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k}\right]$ is the subjective expectation by the behavioral agent, and $\mathbb{E}_t\left[\boldsymbol{X}_{t+k}\right]$ is the rational expectation. The more distant the events in the future, the more the behavioral agent "sees them dimly", i.e. sees them with a dampened cognitive discount factor $\bar{m}^k$ at horizon k (recall that $\bar{m} \in [0,1]$). The parameter $\bar{m}$ models a form of "global cognitive discounting" – discounting future disturbances more (relative to a rational agent) when the forecasting horizon is more distant. Importantly, this implies that all perceived variables will embed some cognitive discounting:[13]

Lemma 2.2 (Cognitive discounting of all variables) *For any variable $z\left(\boldsymbol{X}_t\right)$ with $z\left(0\right) = 0$, the beliefs of the behavioral agent satisfy, for all $k \geq 0$, and linearizing:*

$$\mathbb{E}_t^{BR}\left[z\left(\boldsymbol{X}_{t+k}\right)\right] = \bar{m}^k\mathbb{E}_t\left[z\left(\boldsymbol{X}_{t+k}\right)\right], \tag{11}$$

where $\mathbb{E}_t^{BR}$ is the subjective (behavioral) expectation operator, which uses the misperceived law of motion (8), and $\mathbb{E}_t$ is the rational one, which uses the rational law of motion (5).

---

[10]Section 9 offers a microfoundation for this cognitive discounting assumption.

[11]I particularize the formalism in Gabaix (2016), which is a tractable way to model dynamic programming with limited attention. "Cognitive discounting" was laid out as a possibility in that paper (as a misperception of autocorrelations), but its concrete impact was not studied in any detail there.

[12]When the mean of $\boldsymbol{X}_t$ is not 0, but rather $\boldsymbol{X}_*$ such that $\boldsymbol{X}_* = \boldsymbol{G}\left(\boldsymbol{X}_*, 0\right)$, then the process perceived by the behavioral agent is: $\boldsymbol{X}_{t+1} = \left(1 - \bar{m}\right)\boldsymbol{X}_* + \bar{m}\boldsymbol{G}\left(\boldsymbol{X}_t, \boldsymbol{\epsilon}_{t+1}\right)$. Then, we have, linearizing, $\mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k} - \boldsymbol{X}_*\right] = \bar{m}^k\mathbb{E}_t\left[\boldsymbol{X}_{t+k} - \boldsymbol{X}_*\right]$. In addition, there is no long term growth in this model, as in the basic New Keynesian model. It is easy to introduce it (see Section 12.7 of the Online Appendix). The behavioral agent would be rational with respect to the values around the balanced growth path, but myopic for the deviations from it.

[13]Linearizing, we have $z\left(\boldsymbol{X}\right) = b_{\boldsymbol{X}}^z \boldsymbol{X}$ for some row vector $b_{\boldsymbol{X}}^z$, and:

$$\mathbb{E}_t^{BR}\left[z\left(\boldsymbol{X}_{t+k}\right)\right] = \mathbb{E}_t^{BR}\left[b_{\boldsymbol{X}}^z \boldsymbol{X}_{t+k}\right] = b_{\boldsymbol{X}}^z \mathbb{E}_t^{BR}\left[\boldsymbol{X}_{t+k}\right] = b_{\boldsymbol{X}}^z \bar{m}^k\mathbb{E}_t\left[\boldsymbol{X}_{t+k}\right] = \bar{m}^k\mathbb{E}_t\left[b_{\boldsymbol{X}}^z \boldsymbol{X}_{t+k}\right] = \bar{m}^k\mathbb{E}_t\left[z\left(\boldsymbol{X}_{t+k}\right)\right].$$

For instance, the interest rate perceived in k periods is

$$\mathbb{E}_t^{BR}\left[\bar{r} + \hat{r}\left(\boldsymbol{X}_{t+k}\right)\right] = \bar{r} + \bar{m}^k \mathbb{E}_t\left[\hat{r}\left(\boldsymbol{X}_{t+k}\right)\right].$$

The agent perceives correctly the average interest rate $\bar{r}$ and is globally patient with respect to it, like the rational agent, but he perceives myopically future deviations from the average interest rate (i.e., $\mathbb{E}_t\left[\hat{r}\left(\boldsymbol{X}_{t+k}\right)\right]$ is dampened by $\bar{m}^k$).[14]

2.2 The Firm's Problem

Next, I explore what happens if firms do not fully pay attention to future macro variables either. Firm i sets a price P_{it}. As the final good is produced competitively in quantity $Y_t = \left(\int_0^1 Y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di\right)^{\frac{\varepsilon}{\varepsilon-1}}$, the aggregate price level is:

$$P_t = \left(\int_0^1 P_{it}^{1-\varepsilon} di\right)^{\frac{1}{1-\varepsilon}}. \tag{12}$$

Firms have the usual Calvo pricing friction: at each period, they can reset their price only with probability $1 - \theta$.

Setup: Objective reality and the rational firm's problem   Consider a firm i, and call $q_{i\tau} := \ln \frac{P_{i\tau}}{P_\tau} = p_{i\tau} - p_\tau$ its real log price at time τ. Its real profit is

$$v_\tau = \left(\frac{P_{i\tau}}{P_\tau} - MC_\tau\right)\left(\frac{P_{i\tau}}{P_\tau}\right)^{-\varepsilon} c_\tau.$$

Here $\left(\frac{P_{i\tau}}{P_\tau}\right)^{-\varepsilon} c_\tau$ is the total demand for the firm's good, with c_τ aggregate consumption; $MC_t = (1 - \tau_f)\frac{\omega_t}{e^{\zeta t}} = (1 - \tau_f) e^{-\mu_t}$ is the real marginal cost; $\mu_t := \zeta_t - \ln \omega_t$ is the labor wedge, which is zero at efficiency. A corrective wage subsidy $\tau_f = \frac{1}{\varepsilon}$ ensures that there are no price distortions on average. For simplicity I assume that this subsidy is financed by a lump-sum tax on firms, which affects v_τ by an additive value, so that it does not change the pricing decision: v_τ is the firm's profit before the lump-sum tax. It is equal to:

$$v^0\left(q_{i\tau}, \mu_\tau, c_\tau\right) := \left(e^{q_{i\tau}} - (1 - \tau_f) e^{-\mu_\tau}\right) e^{-\varepsilon q_{i\tau}} c_\tau. \tag{13}$$

I fix a date t and consider the worldview of a firm simulating future dates $\tau \geq t$. Call $\boldsymbol{X}_\tau$ the extended macro state vector $\boldsymbol{X}_\tau = \left(\boldsymbol{X}_\tau^{\mathcal{M}}, \Pi_\tau\right)$ where $\Pi_\tau := p_\tau - p_t = \pi_{t+1} + \cdots + \pi_\tau$ is inflation between times t and τ, and $\boldsymbol{X}_\tau^{\mathcal{M}}$ is the vector of macro variables, including TFP ζ_t as well as possible announcements about future policy. Then, if the firm hasn't changed its price between t and τ, its real price is $q_{i\tau} = q_{it} - \Pi_\tau$, so the flow profit at τ is:

$$v\left(q_{it}, \boldsymbol{X}_\tau\right) := v^0\left(q_{it} - \Pi\left(\boldsymbol{X}_\tau\right), \mu\left(\boldsymbol{X}_\tau\right), c\left(\boldsymbol{X}_\tau\right)\right), \tag{14}$$

where $\Pi_\tau := \Pi\left(\boldsymbol{X}_\tau\right)$ is aggregate future inflation, and similarly for μ and c. A traditional Calvo firm

---

[14]Here the agent anchors variables on their steady state values. One might imagine anchoring variables on their present value, which leads to interesting variants (see Section 12.3), in particular leading to a high marginal propensity to consume.

which can reset its price at t wants to choose the optimal real price q_{it} to maximize total profits, as in:

$$\max_{q_{it}} \mathbb{E}_t \sum_{\tau=t}^{\infty} (\beta\theta)^{\tau-t} \frac{c(\boldsymbol{X}_\tau)^{-\gamma}}{c(\boldsymbol{X}_t)^{-\gamma}} v(q_{it}, \boldsymbol{X}_\tau), \tag{15}$$

where $\frac{c(\boldsymbol{X}_\tau)^{-\gamma}}{c(\boldsymbol{X}_t)^{-\gamma}}$ is the adjustment in the stochastic discount factor between t and τ.

Setup: Reality perceived by a behavioral firm   The behavioral firm faces the same problem, with a less accurate view of reality. Most importantly, I posit that the behavioral firm also perceives the future via the cognitive discounting mechanism in (8). The behavioral firm wants to optimize its initial real price level q_{it} under the perceived law of motion given in (8), reflecting cognitive dampening. That is, the behavioral firm solves:[15]

$$\max_{q_{it}} \mathbb{E}_t^{BR} \sum_{\tau=t}^{\infty} (\beta\theta)^{\tau-t} \frac{c(\boldsymbol{X}_\tau)^{-\gamma}}{c(\boldsymbol{X}_t)^{-\gamma}} v(q_{it}, \boldsymbol{X}_\tau). \tag{16}$$

Linearizing around the deterministic steady state, $\frac{c(\boldsymbol{X}_\tau)^{-\gamma}}{c(\boldsymbol{X}_t)^{-\gamma}} \simeq 1$, so that term will not matter in the linearizations.

2.3   Model Solution

We have laid out the model's assumptions. We now proceed to solving it.

Household optimization   We can now study the household's optimization problem under both the rational and the behavioral setup. The Euler equation of a rational agent is: $\mathbb{E}_t \left[\beta R_t \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \right] = 1$. Linearizing, we get:[16]

$$\hat{c}_t = \mathbb{E}_t [\hat{c}_{t+1}] - \frac{1}{\gamma R} \hat{r}_t. \tag{17}$$

This is the traditional derivation of the investment-savings (IS) curve, with rational agents.

Now call $c(\boldsymbol{X}_t, k_t)$ the equilibrium consumption of the behavioral agent. Under the agent's subjective model, we have:[17] $\mathbb{E}_t^{BR}[\beta R_t \left(\frac{c(\boldsymbol{X}_{t+1}, k_{t+1})}{c(\boldsymbol{X}_t, k_t)} \right)^{-\gamma}] = 1$. Now, in general equilibrium, income and consumption are the same (so $c_t = \bar{y} + \hat{y}(N_t, \boldsymbol{X}_t)$) and private wealth is $k_t = 0$. Hence, given (6), the agent correctly

---

[15]Here I use the same $\bar{m}$ for consumer and firms. If firms had their own rate of cognitive discounting $\bar{m}^f$, then one would simply replace $\bar{m}$ by $\bar{m}^f$ in the expression for M^f and in (26).

[16]Indeed, using $\beta R = 1$ and $R_t = R + \hat{r}_t$,

$$1 = \mathbb{E}_t \left[\beta R_t \left(\frac{c_{t+1}}{c_t} \right)^{-\gamma} \right] = \mathbb{E}_t \left[\beta R \left(1 + \frac{\hat{r}_t}{R} \right) \left(\frac{1+\hat{c}_{t+1}}{1+\hat{c}_t} \right)^{-\gamma} \right] \simeq 1 + \frac{\hat{r}_t}{R} - \gamma \mathbb{E}_t[\hat{c}_{t+1} - \hat{c}_t],$$

which gives (17). Galí (2015) does not have the $\frac{1}{R}$ term as he defines the interest rate as $r_t^{\text{Galí}} := \ln R_t$, whereas in the present paper it is defined as $r_t := R_t - 1$, so that $\hat{r}_t^{\text{Galí}} \equiv \frac{\hat{r}_t}{R}$. The predictions are the same, adjusting for the slightly different convention.

[17]At time t, the agent has an expectation about the distribution of his future consumption $c(\boldsymbol{X}_{t+1}, k_{t+1})$, but does not commit to that consumption. Indeed, he'll pick his consumption c_{t+1} only at time $t+1$, as more news (rationally) or a better perception of the world (behaviorally) is revealed at time $t+1$.

anticipates that her beginning of period $t+1$ private wealth will be $k_{t+1} = 0$.[18] It follows that aggregate consumption $c(\boldsymbol{X}_t) = c(\boldsymbol{X}_t, 0)$ satisfies

$$\mathbb{E}_t^{BR}\left[\beta R_t \left(\frac{c(\boldsymbol{X}_{t+1})}{c(\boldsymbol{X}_t)}\right)^{-\gamma}\right] = 1.$$

Linearizing, this gives:

$$\hat{c}(\boldsymbol{X}_t) = \mathbb{E}_t^{BR}[\hat{c}(\boldsymbol{X}_{t+1})] - \frac{1}{\gamma R}\hat{r}_t.$$

Now, by Lemma 2.2, $\mathbb{E}_t^{BR}[\hat{c}_t(\boldsymbol{X}_{t+1})] = \bar{m}\mathbb{E}_t[\hat{c}_t(\boldsymbol{X}_{t+1})]$, so we obtain

$$\hat{c}_t = M\mathbb{E}_t[\hat{c}_{t+1}] - \sigma\hat{r}_t, \tag{18}$$

with $M = \bar{m}$ and $\sigma = \frac{1}{\gamma R}$. Equation (18) is a "discounted aggregate Euler equation". I call M the macro parameter of attention. Here $M = \bar{m}$, but in more general specifications coming later, $M \neq \bar{m}$, so, anticipating them, I keep the notation M for macro attention.

The behavioral IS curve   We can now derive the IS curve for the behavioral household. Let us link (18) to the output gap. First, the static first order condition for labor supply holds:

$$N_t^\phi = \omega_t c_t^{-\gamma}. \tag{19}$$

Next, call c_t^n and r_t^n the natural rate of output and interest, defined as the quantity of output and interest that would prevail if we removed all pricing frictions, and use hats to denote them as deviations from the steady state, $\hat{c}_t^n := c_t^n - \bar{c}$ and $\hat{r}_t^n := r_t^n - \bar{r}$. The natural rate of output is easy to derive;[19] it is

$$\hat{c}_t^n = \frac{1+\phi}{\gamma+\phi}\zeta_t. \tag{20}$$

Next, note that equation (18) also holds in that "natural" economy that would have no pricing frictions. So,

$$\hat{c}_t^n = M\mathbb{E}_t[\hat{c}_{t+1}^n] - \sigma\hat{r}_t^n, \tag{21}$$

which gives that the natural rate of interest $r_t^n = r_t^{n0}$, where

$$r_t^{n0} = \bar{r} + \frac{1+\phi}{\sigma(\gamma+\phi)}(M\mathbb{E}_t[\zeta_{t+1}] - \zeta_t). \tag{22}$$

I call this interest rate r_t^{n0} the "pure" natural rate of interest—this is the interest rate that prevails in an economy without pricing frictions, and undisturbed by government policy (in particular, budget deficits). So when there are no budget deficits, as is the case here, we have that $r_t^n = r_t^{n0}$ — but in later

---

[18]When the agent has non-zero private wealth (which is the case with taxes) or when she misperceives her income, the derivation is more complex, as we shall see in Section 6.1.

[19]With flexible prices, the aggregate resource constraint is $c_t = e^{\zeta_t}N_t$, and the wage is $\omega_t = e^{\zeta_t}$. Together with (19), we obtain the natural rate of output, $\ln c_t^n = \frac{1+\phi}{\gamma+\phi}\zeta_t$; linearizing around $\bar{c} = 1$, so that $\ln c_t^n \simeq \hat{c}_t^n$, we get the announced value.

11

extensions the two concepts will differ. Behavioral forces don't change the natural rate of output, but they do change the pure natural rate of interest.[20]

The output gap is $x_t := \hat{c}_t - \hat{c}_t^n$. Then, taking (18) minus (21), we obtain:

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] - \sigma\left(\hat{r}_t - \hat{r}_t^n\right),\tag{23}$$

with $M = \bar{m}$. Rearranging, $\hat{r}_t - \hat{r}_t^n = (r_t - \bar{r}) - (r_t^n - \bar{r}) = i_t - \mathbb{E}_t\left[\pi_{t+1}\right] - r_t^n$, where i_t is the nominal interest rate. We obtain the following result:[21]

Proposition 2.3 (Behavioral discounted Euler equation) *In equilibrium, the output gap x_t follows:*

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] - \sigma\left(i_t - \mathbb{E}_t\left[\pi_{t+1}\right] - r_t^n\right),\tag{24}$$

where $M = \bar{m} \in [0,1]$ is the macro attention parameter, and $\sigma := \frac{1}{\gamma R}$. In the rational model, $M = 1$.

The behavioral NK IS curve (23) implies:

$$x_t = -\sigma\sum_{k\geq 0} M^k\mathbb{E}_t\left[\hat{r}_{t+k} - \hat{r}_{t+k}^n\right].\tag{25}$$

In the rational case with $M = 1$, a one-period change in the real interest rate $\hat{r}_{t+k}$ in 1000 periods has the same impact on the output gap as a change occurring today. This is intuitively very odd, and is an expression of the forward guidance puzzle. However, when $M < 1$, a change occurring in 1000 periods has a much smaller impact than a change occurring today.[22]

Firm optimization   Given the objective function (16), we can study the solution to the behavioral firm's problem. The nominal price that firm i will choose will be $p_t^* = q_{it} + p_t$, and its value is as in the following lemma (the derivation is in Section 11.2).

Lemma 2.4 (Optimal price for a behavioral firm resetting its price) *A behavioral firm resetting its price at time t will set it to a value p_t^* equal to:*

$$p_t^* = p_t + (1 - \beta\theta)\sum_{k=0}^{\infty}(\beta\theta\bar{m})^k\,\mathbb{E}_t\left[\pi_{t+1} + \ldots + \pi_{t+k} - \mu_{t+k}\right],\tag{26}$$

where $\bar{m}$ is the overall cognitive discounting factor.

The Appendix (Section 11.2) traces out the macroeconomic implications of the producer price-setting policy (26), yielding the behavioral New Keynesian Phillips curve, which I present in Proposition 2.5.

---

[20]In a model with physical capital, behavioral forces would change the natural rate of output.

[21]Substantially, the agent anchors on the steady state. This implies that cognitive discounting is about the deviation of output from the steady state (and not just the output gap).

[22]To think about finance, the model would allow for a fringe of rational financial arbitrageurs (with vanishingly small consumption share), so that then all assets would be priced rationally. The present model generates no risk premia since all risks are small. I defer to future research the exploration of asset pricing implications of this sort of model, which would require adding non-trivial risks, e.g. disaster risk.

2.4 A Behavioral New Keynesian Model

Combining the results of the household's and the firm's problems in general equilibrium, we obtain the following two-equation synthesis. This two-equation model features behavioral versions of both the dynamic IS curve and the New Keynesian Phillips curve:

Proposition 2.5 (Behavioral New Keynesian model – two equation version) *We have the following behavioral version of the New Keynesian model, for the behavior of the output gap x_t and inflation π_t:*

$$x_t = M\mathbb{E}_t[x_{t+1}] - \sigma(i_t - \mathbb{E}_t\pi_{t+1} - r_t^n) \quad \text{(IS curve)}, \tag{27}$$

$$\pi_t = \beta M^f \mathbb{E}_t[\pi_{t+1}] + \kappa x_t \quad \text{(Phillips curve)}, \tag{28}$$

where $M, M^f \in [0,1]$ are the aggregate-level attention parameters of consumers and firms, respectively:

$$M = \bar{m}, \qquad \sigma = \frac{1}{\gamma R}, \tag{29}$$

$$M^f = \bar{m}\left(\theta + \frac{1 - \beta\theta}{1 - \beta\theta\bar{m}}(1 - \theta)\right). \tag{30}$$

In this baseline specification, which abstracts from fiscal policy, r_t^n corresponds to the "pure" natural interest rate r_t^{n0} that prevails with zero deficits, derived in (22). In the traditional model, $\bar{m} = 1$, so that $M = M^f = 1$. In addition, the Phillips curve slope is $\kappa = \bar{\kappa}$, where $\bar{\kappa} = \left(\frac{1}{\theta} - 1\right)(1 - \beta\theta)(\gamma + \phi)$ is the slope obtained with fully rational firms.[23]

Like consumers, firms can be fully attentive to all idiosyncratic terms (something that would be easy to include), such as the idiosyncratic part of their productivity. For the purposes of this result, they simply have to pay limited attention to macro outcomes. If we include idiosyncratic terms, and firms are fully attentive to them, the aggregate NK Phillips curve does not change. Also, firms and consumers are still fully rational for steady state variables (e.g., in the steady state firms discount future profits at rate $R = \frac{1}{\beta}$). It is only their sensitivity to deviations from the deterministic steady state that is partially myopic. Under a realistic calibration, aggregate inflation is more forward-looking (M^f is higher) when prices are sticky for a longer period of time (θ is higher) and when firms are more attentive to future macroeconomic outcomes ($\bar{m}$ is higher).[24] In the traditional model the coefficient on future inflation in (28) is exactly β and, quite miraculously, does not depend on the adjustment rate of prices θ. In the behavioral model, in contrast, the coefficient (βM^f) is higher when prices are stickier for longer (higher θ).

2.5 Values Used in the Numerical Examples

Table 1 summarizes the main sufficient statistics for the output of the model, summarized in Figures 1–5. These values can in turn be rationalized in terms of "ancillary" parameters shown in Table 2. I call these parameters "ancillary" because they matter only via their impact on the aforementioned sufficient statistics listed in Table 1. For instance, the value of $\kappa = \left(\frac{1}{\theta} - 1\right)(1 - \beta\theta)(\gamma + \phi)$ can come

---

[23]While in this baseline formulation we have that $\kappa = \bar{\kappa}$, this will not be true in the enriched model that will be presented in Section 6.2. For consistency of presentation throughout the paper, I therefore keep these two notations separate.

[24]Note that $\beta R = 1$ pins down the value of β. So, one could not accommodate an anomalous Phillips curve by just changing β: that would automatically change the steady state interest rate.

from many combinations of θ, γ, ϕ etc. Table 2 shows one such combination. The values are broadly consistent with those of the New Keynesian literature.[25]

The inattention parameters are chosen to be close to the myopia found in Galí and Gertler (1999) and Lindé (2005) for the Phillips curve, and Fuhrer and Rudebusch (2004) for the IS curve.[26] In particular, Galí and Gertler (1999) find that we need $\beta M^f \simeq 0.8$ at the quarterly frequency; given that $\beta \simeq 1$, that leads to an attention parameter of $M^f \simeq 0.8$. Fuhrer and Rudebusch (2004) estimate an IS curve and find $M \simeq 0.65$, with a standard error of 0.15. Relatedly, the literature on the forward guidance puzzle concludes, qualitatively, that $M < 1$. Since a value $M \simeq 0.65$ is quite extreme, here I adopt a more conservative value of $M = 0.85$.

For intuition, if the time period is taken to be a quarter, the calibrated macro attention parameter $M = 0.85$ implies that the half-life of consumer attention is roughly over one year — in the sense that the consumer pays just over half as much attention to an innovation that is to come a year into the future relative to the attention she pays to an innovation today (since $M^4 \simeq 0.5$).[27]

Table 1: Key Parameter Inputs

Cognitive discounting by consumers and firms	$M = 0.85,\ M^f = 0.80$
Sensitivity to interest rates	$\sigma = 0.20$
Slope of the Phillips curve	$\kappa = 0.11$
Rate of time preference	$\beta = 0.99$
Relative welfare weight on output	$\vartheta = 0.02$

Notes. This table reports the coefficients used in the model. Units are quarterly.

[25]The value of θ is admittedly quite high, as in this baseline model it helps match a low value of the Phillips curve slope κ. Other mechanisms that are consistent with lower levels of price stickiness θ can also generate a low slope of the Phillips curve (using firms' inattention to macro conditions). See the discussion in Section 6.2 and footnote 70.

[26]In general, it is quite hard to identify M, M^f in aggregate data separately from other model parameters. For illustration, consider a simple model specification with fully rigid prices ($\pi_t = \kappa = 0$) and in which $i_t = r_t^n$ always. The econometric version of the IS curve is $x_t = M\mathbb{E}[x_{t+1}] + \eta_t$, where η_t is an exogenous shock that is not directly observed. Suppose η_t follows an AR(1) process with innovation variance σ_η^2 and persistence ρ. The IS curve dynamics then imply that $x_t = \frac{1}{1-M\rho}\eta_t$. A simple time-series regression of x_t on x_{t-1} identifies ρ as well as $\frac{\sigma_\eta}{1-M\rho}$, which is the standard deviation of the innovations to x_t. Hence M and σ_η are not separately identified from the time series of x_t. In general, New Keynesian parameters are also hard to estimate (Mavroeidis et al. (2014)).

[27]This literature provides evidence on the macro parameters of attention M, M^f. The 2018 NBER working paper version of the present paper provides a tentative Bayesian estimation of the entire model that finds estimated parameters consistent with those in Table 1. The estimation in that draft should be taken as preliminary at best since further well-identified empirical work will be necessary in order to reach definitive conclusions. Gathering evidence on micro parameters $\bar{m}$ would be much more costly. However, using micro data, Ganong and Noel (2019) find evidence for micro-level cognitive discounting, so that progress is being made in that direction too. This paper provides evidence that explanations for the failure of the undiscounted Euler equation that are based on financial frictions (for example, credit constraints as in Kaplan et al. (2018)) should be complemented by behavioral ones.

Table 2: ANCILLARY PARAMETERS

Coefficient of risk aversion	$\gamma = 5$
Inverse of Frisch elasticity	$\phi = 1$
Survival rates of prices	$\theta = 0.875$
Demand elasticity	$\varepsilon = 5.3$
Cognitive discounting	$\bar{m} = 0.85$

Notes. This table reports the coefficients used in the model to generate the parameters of Table 1. Units are quarterly.

3 Consequences

Here I explore the consequences of the model. They mostly depend on the reduced form of Proposition 2.5, rather than its detailed microfoundations. Hence, they would apply to other models when they have a similar reduced form (for example, Woodford (2018) and Angeletos and Lian (2018)).

3.1 The Taylor Principle Reconsidered: Equilibria are Determinate Even with a Fixed Interest Rate

The traditional model suffers from the existence of a continuum of multiple equilibria when monetary policy is passive. We will now see that if consumers are boundedly rational enough, there is just one unique (bounded) equilibrium. As monetary policy is passive at the ZLB, this topic will have strong impacts for the economy at the ZLB.

I assume that the central bank sets the nominal interest i_t in a Taylor rule fashion:

$$i_t = \phi_\pi \pi_t + \phi_x x_t + j_t, \tag{31}$$

where j_t is typically just a constant.[28] Calculations show that the system of Proposition 2.5 can be represented as:[29]

$$\boldsymbol{z}_t = \boldsymbol{A} \mathbb{E}_t [\boldsymbol{z}_{t+1}] + \boldsymbol{b} a_t, \tag{32}$$

where

$$\boldsymbol{z}_t := (x_t, \pi_t)'$$

will be called the state vector,[30] $a_t := j_t - r_t^n$ (as in "action") is the baseline tightness of monetary

---

[28]The reader will want to keep in mind the case of a constant $j_t = \bar{j}$. Generally, j_t is a function $j_t = j(\boldsymbol{X}_t)$ where $\boldsymbol{X}_t$ is a vector of primitives that are not affected by (x_t, π_t), e.g. the natural rate of interest coming from stochastic preferences and technology.

[29]It is easier (especially for higher-dimensional variants) to proceed with the matrix $\boldsymbol{B} := \boldsymbol{A}^{-1}$, write the system as $\mathbb{E}_t [\boldsymbol{z}_{t+1}] = \boldsymbol{B}\boldsymbol{z}_t + \widetilde{\boldsymbol{b}} a_t$, and to reason on the eigenvalues of $\boldsymbol{B}$:

$$\boldsymbol{B} = \frac{1}{M\beta^f} \begin{pmatrix} \beta^f (1 + \sigma\phi_x) + \kappa\sigma & -\sigma \left(1 - \beta^f \phi_\pi\right) \\ -\kappa M & M \end{pmatrix}.$$

[30]I call $\boldsymbol{z}_t$ the "state vector" with some mild abuse of language. It is an outcome of the deeper state vector $\boldsymbol{X}_t$.

policy (if the government pursues the first best, $a_t = 0$), $\boldsymbol{b} = \frac{-\sigma}{1+\sigma(\phi_x + \kappa\phi_\pi)} (1, \kappa)'$ and

$$\boldsymbol{A} = \frac{1}{1 + \sigma\left(\phi_x + \kappa\phi_\pi\right)} \begin{pmatrix} M & \sigma\left(1 - \beta^f \phi_\pi\right) \\ \kappa M & \beta^f\left(1 + \sigma\phi_x\right) + \kappa\sigma \end{pmatrix}, \tag{33}$$

where I use the notation

$$\beta^f := \beta M^f. \tag{34}$$

The next proposition generalizes the well-known Taylor determinacy criterion to behavioral agents. I assume that ϕ_π and ϕ_x are weakly positive (the proof indicates a more general criterion).

Proposition 3.1 (Equilibrium determinacy with behavioral agents) *There is a determinate equilibrium (i.e., all of $\boldsymbol{A}$'s eigenvalues are less than 1 in modulus) if and only if:*

$$\phi_\pi + \frac{\left(1 - \beta M^f\right)}{\kappa}\phi_x + \frac{\left(1 - \beta M^f\right)\left(1 - M\right)}{\kappa\sigma} > 1. \tag{35}$$

In particular, when monetary policy is passive (i.e., when $\phi_\pi = \phi_x = 0$), we have a determinate equilibrium if and only if bounded rationality is strong enough, in the sense that

$$\text{Strong enough bounded rationality condition: } \frac{\left(1 - \beta M^f\right)\left(1 - M\right)}{\kappa\sigma} > 1. \tag{36}$$

Condition (36) does not hold in the traditional model, where $M = 1$. The condition means that agents are boundedly rational enough (i.e., M is sufficiently less than 1) and the firm-level pricing or cognitive frictions are large enough.[31] Quantitatively, it is quite easy to satisfy this criterion.[32]

Why does bounded rationality eliminate multiple equilibria? This is because boundedly rational agents are less reactive to the future, hence less reactive to future agents' decisions. Therefore, bounded rationality lowers the complementarity between agents' actions (their consumptions). That force dampens the possibility of multiple equilibria.[33,34]

Condition (36) implies that the two eigenvalues of $\boldsymbol{A}$ are less than 1. This implies that the equilibrium is determinate.[35] This is different from the traditional NK model, in which there is a continuum of non-explosive monetary equilibria, given that one root is greater than 1 (as condition (36) is violated in the traditional model).

This absence of multiple equilibria is important, in particular when the central bank keeps an interest peg (e.g. at 0% because of the ZLB).

---

[31]Bilbiie (2018) independently derives a condition similar to equation (35) with $M^f = 1$, $\phi_x = 0$, and in the context of a heterogenous-agents New Keynesian model .

[32]Call $g(M) = \frac{(1-M)(1-\beta M)}{\kappa\sigma} - 1$ (to simplify this discussion, I take $M = M^f$). The behavioral Taylor criterion (35) is that $g(M) > 0$, i.e. $M < M^*$ where $g(M^*) = 0$. Using the calibration, this is the case if and only if $M^* \simeq 0.86$. If we divide $\kappa\sigma$ by 10 (which is not difficult, given the small values of κ and σ often estimated) we get $M^* \simeq 0.95$.

[33]This theme that bounded rationality reduces the scope for multiple equilibria is general, and also holds in simple static models. I plan to develop it separately.

[34]One could also introduce nominal illusion as consumers perceiving the inflation to be $\pi^{BR}(\boldsymbol{X}_t) = m_\pi^c \pi(\boldsymbol{X}_t)$, where $m_\pi^c \in [0,1]$ parameterizes limited attention to inflation. In the IS curve (27), that will lead to replacing $\mathbb{E}_t \pi_{t+1}$ by $m_\pi^c \mathbb{E}_t \pi_{t+1}$. More surprisingly, the Taylor criterion is modified by replacing, in the right-hand side of (35), the 1 by m_π^c (see Section 12.6). Again, bounded rationality makes the Taylor criterion easier to satisfy.

[35]The condition does not prevent unbounded or explosive equilibria, the kind that Cochrane (2011) analyzes. My take is that this issue is interesting (as are rational bubbles in general), but that the main practical problem is to eliminate bounded equilibria. The present behavioral model does that well.

Permanent interest rate peg. First, take the (admittedly extreme) case of a permanent peg. Then, in the traditional model, there is always a continuum of bounded equilibria, technically, because the matrix $\boldsymbol{A}$ has a root greater than 1 (in modulus) when $M = 1$. As a result, there is no definite answer to the question "What happens if the central bank raises the interest rate?" – as one needs to select a particular equilibrium. In this paper's behavioral model, however, we do get a definite non-explosive equilibrium. Indeed, as all roots are less than 1 in modulus, we can simply write:

$$\boldsymbol{z}_t = \mathbb{E}_t \left[\sum_{\tau \geq t} \boldsymbol{A}^{\tau - t} \boldsymbol{b} a_\tau \right]. \tag{37}$$

Cochrane (2017) made the point that we'd expect an economy such as Japan's to be quite volatile, if the ZLB is expected to last forever: conceivably, the economy could jump from one equilibrium to the next at each period. This is a problem for the rational model, which is solved if agents are behavioral enough, so that (36) holds.

Long-lasting interest rate peg. Second, the economy is still very volatile (in the rational model) in the less extreme case of a peg lasting for a long but finite duration. To see this, suppose that the ZLB is expected to last for T periods. Call $\boldsymbol{A}_{ZLB}$ the value of matrix $\boldsymbol{A}$ in (33) when $\phi_\pi = \phi_x = j = 0$ in the Taylor rule. Then, the system (32) is, at the ZLB ($t \leq T$): $\boldsymbol{z}_t = \mathbb{E}_t \boldsymbol{A}_{ZLB} \boldsymbol{z}_{t+1} + \underline{\boldsymbol{b}}$ with $\underline{\boldsymbol{b}} := (1, \kappa) \sigma \underline{r}$, where $\underline{r} < 0$ is the real interest rate that prevails during the ZLB. Iterating forward, we have:

$$\boldsymbol{z}_0 (T) = \left(\boldsymbol{I} + \boldsymbol{A}_{ZLB} + ... + \boldsymbol{A}_{ZLB}^{T-1} \right) \underline{\boldsymbol{b}} + \boldsymbol{A}_{ZLB}^T \mathbb{E}_0 \left[\boldsymbol{z}_T \right]. \tag{38}$$

Here I note $\boldsymbol{z}_0 (T)$, the value of the state at time 0, given the ZLB will last for T periods. Let us focus on the last term, $\boldsymbol{A}_{ZLB}^T \mathbb{E}_0 [\boldsymbol{z}_T]$. In the traditional case, one of the eigenvalues of $\boldsymbol{A}_{ZLB}$ is greater than 1 in modulus. This implies that very small changes to today's expectations of economic conditions after the ZLB (i.e., to $\mathbb{E}_0 [\boldsymbol{z}_T]$), have an unboundedly large impact today ($\lim_{T \to \infty} \| \boldsymbol{A}_{ZLB}^T \| = \infty$). Hence, we would expect the economy to be very volatile today, provided the ZLB period is long though finite, and a reasonable amount of fluctuating uncertainty about future policy.

3.2 The ZLB is Less Costly with Behavioral Agents

What happens when economies are at the ZLB? The rational model makes very stark predictions, which this behavioral model overturns. To see this, I follow the thought experiment in Werning (2012) (building on Eggertsson and Woodford (2003)), but with behavioral agents. I take $r_t^n = \underline{r}$ for $t \leq T$, and $r_t^n = \bar{r}$ for $t > T$, with $\underline{r} < 0 \leq \bar{r}$. I assume that for $t > T$, the central bank implements $x_t = \pi_t = 0$ by setting $i_t = \bar{r} + \phi_\pi \pi_t$ with $\phi_\pi > 1$, so that in equilibrium $i_t = \bar{r}$. At time $t < T$, I suppose that the economy is at the ZLB, so that $i_t = 0$.

Proposition 3.2 *Call $x_0 (T)$ the output gap at time 0, given the ZLB will lasts for T periods. In the traditional rational case, we obtain an unboundedly intense recession as the length of the ZLB increases: $\lim_{T \to \infty} x_0 (T) = -\infty$. This also holds when myopia is mild, i.e. (36) fails. However, suppose cognitive myopia is strong enough, i.e. (36) holds. Then, we obtain a boundedly intense recession:*

$$\lim_{T \to \infty} x_0 (T) = \frac{\sigma(1 - \beta M^f)}{(1 - M)(1 - \beta M^f) - \kappa \sigma} \underline{r} < 0. \tag{39}$$
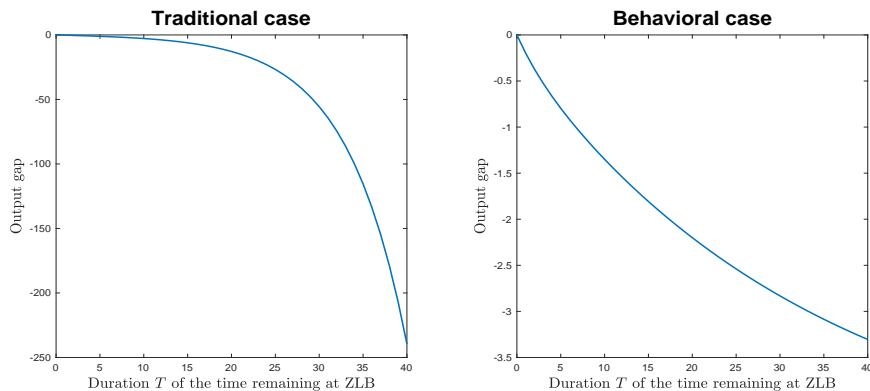
17

Traditional case / **Behavioral case**

Figure 1: This figure shows the output gap $x_0(T)$ at time 0, given that the economy will be at the ZLB for T more periods. The left panel is the traditional New Keynesian model, the right panel the present behavioral model. Parameters are the same in both models, except for the attention parameters M, M^f which are equal to 1 in the rational model. The natural rate at the ZLB is -1%. Output gap units are percentage point. Time units are quarters.

We see how impactful myopia can be.[36] Myopia has to be stronger when agents are highly sensitive to the interest rate (high σ) and price flexibility is high (high κ). High price flexibility makes the system very reactive, and a high myopia is useful to counterbalance that.[37] Figure 1 shows the result. The left panel shows the traditional model, while the right one shows the behavioral model. The parameters are the same in both models, except that attention is lower in the behavioral model. In the left panel, we see how costly the ZLB is: mathematically it is unboundedly costly as it becomes more long-lasting, displaying an exponentially bad recession as the ZLB is more long-lasting. In contrast, in this behavioral model, in the right panel we see a finite, though prolonged cost. Reality looks more like the mild slump of the behavioral model (right panel) – something like Japan since the 1990s – rather than the frightful abyss of the rational model (left panel), which is something like Japan in 1946. This sort of effect could be useful to empirically show that (36) likely holds.[38]

3.3   Forward Guidance is Much Less Powerful

Suppose that the central bank announces at time 0 that in T periods it will perform a one-period, 1 percent real interest rate cut. What is the impact on today's inflation? This is the thought experiment analyzed by McKay et al. (2016) with rational agents, which I pursue here with behavioral agents.

Figure 2 illustrates the effect. In the left panel, the whole economy is rational. We see that the further away the policy, the bigger the impact today – this is quite surprising, hence the term "forward guidance puzzle". In the middle panel, consumers are behavioral but firms are rational, while in the right panel both consumers and firms are behavioral. We see that indeed, announcements about very

---

[36]McKay et al. (2017) recognize the point that a discounted aggregate Euler equation leads to a shallower ZLB recession, but they operate entirely within the "mild myopia" case in the terminology of Proposition 3.2 (in which the recession's intensity is reduced but remains unbounded), and do not detect the existence of the "strong myopia" case.

[37]The "paradox of flexibility" still holds though in a dampened way: if prices are more flexible, κ is higher (Proposition 2.5), and the higher disinflation worsens the recession at the ZLB (Eggertsson and Krugman (2012), Werning (2012)). However, bounded rationality moderates this, by lowering κ and M in (39).

[38]In this case, the economy is better off if agents are not too rational. This quite radical change of behavior is likely to hold in other contexts. For instance, in those studied by Kocherlakota (2016) where the very long run matters a great deal, it is likely that a modicum of bounded rationality would change the behavior of the economy considerably.
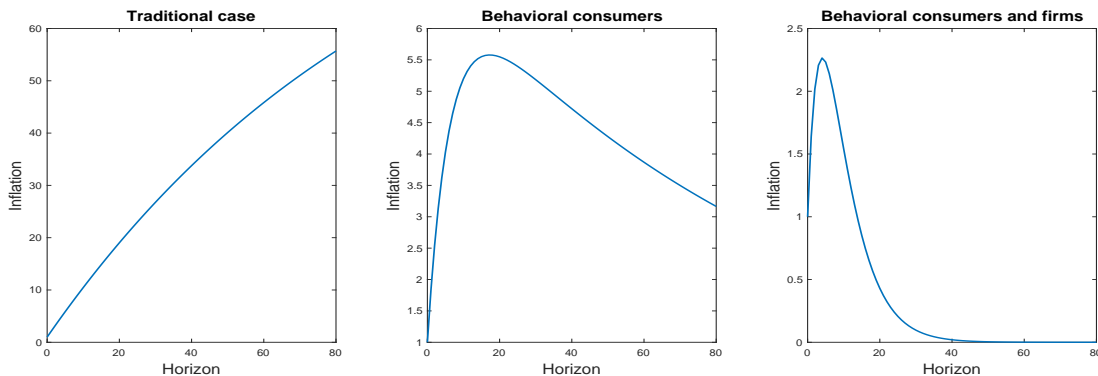
Figure 2: This Figure shows the response of current inflation to forward guidance about a one-period interest rate cut in T quarters, compared to an immediate rate change of the same magnitude. Left panel: traditional New Keynesian model. Middle panel: model with behavioral consumers and rational firms. Right panel: model with behavioral consumers and firms. Parameters are the same in both models, except for the attention parameters M, M^f which are equal to 1 in the rational model.

distant policy changes have vanishingly small effects with behavioral agents – but they have the biggest effect with rational agents.[39,40]

4 Optimal Monetary Policy

4.1 Welfare with Behavioral Agents and the Central Bank's Objective

Optimal policy needs a welfare criterion. The task is simplified by the fact that we have a representative agent. Welfare here is taken to be the expected utility of the representative agent, $\widetilde{W} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t, N_t)$, under the objective expectation. This is as in much of behavioral economics, which views behavioral agents as using heuristics, but experience utility from consumption and leisure like rational agents.[41] Hence, the welfare criterion in this model is the same as in the traditional Woodford (2003b) formulation. I express $\widetilde{W} = W^* + W$, where W^* is first best welfare, and W is the deviation from the first best. The following lemma derives it.

---

[39]Formally, we have $x_t = M x_{t+1} - \sigma \hat{r}_t$, with $\hat{r}_T = -\delta = -1\%$ and $\hat{r}_t = 0$ if $t \neq T$. So $x_t = \sigma M^{T-t} \delta$ for $t \leq T$ and $x_t = 0$ for $t > T$. This implies that the time-0 response of inflation to a one-period interest-rate cut T periods into the future is:

$$\pi_0(T) = \kappa \sum_{t \geq 0} \left(\beta^f\right)^t x_t = \kappa \sigma \sum_{t=0}^{T} \left(\beta^f\right)^t M^{T-t} \delta = \kappa \sigma \frac{M^{T+1} - \left(\beta^f\right)^{T+1}}{M - \beta^f} \delta.$$

A rate cut in the very distant future has a powerful impact on today's inflation ($\lim_{T \to \infty} \pi_0(T) = \frac{\kappa \sigma}{1 - \beta^f} \delta$) in the rational model ($M = 1$), and no impact at all in the behavioral model ($\lim_{T \to \infty} \pi_0(T) = 0$ if $M < 1$)

[40]When attention is endogenous, the analysis could become more subtle. Indeed, if other agents are more attentive to the forward announcement by the Fed, their impact will be bigger, and a consumer will want to be more attentive to it. This positive complementarity in attention could create multiple equilibria in effective attention M. I do not pursue that here.

[41]In particular I use the objective (not subjective) expectations (so that in (40) the discount rate is β, not e.g. βM). Also, I do not include thinking costs in the welfare metric. One reason is that thinking costs are very hard to measure (revealed preference arguments apply only if attention is exactly optimally set, something which is controversial). In the terminology of Farhi and Gabaix (2017), we are in the "no attention in welfare" case. Angeletos and La'O (2008, 2018) and Paciello and Wiederholt (2013) study policy when the planner "internalizes" the cognitive constraints of the agents.

Lemma 4.1 (Welfare) *The welfare loss from inflation and output gap is*

$$W = -K\mathbb{E}_0 \sum_{t=0}^{\infty} \frac{1}{2}\beta^t \left(\pi_t^2 + \vartheta x_t^2\right) + W_-, \tag{40}$$

where $\vartheta = \frac{\bar{\kappa}}{\varepsilon}$, $K = u_c c\left(\gamma + \phi\right)\frac{\varepsilon}{\bar{\kappa}}$, and W_- is a constant (made explicit in (208)), $\bar{\kappa}$ is the Phillips curve coefficient with rational firms (derived in Proposition 2.5), and ε is the elasticity of demand.

4.2 Optimal Policy with no ZLB Constraints: Response to Changes in the Natural Interest Rate

Suppose that there are productivity or discount factor shocks (the latter are not explicitly in the basic model, but can be introduced in a straightforward way). This changes the natural real interest rate, $r_t^n = r_t^{n0}$. To find the policy ensuring the first best (a zero output gap and inflation), we inspect the two equations of this behavioral model, (27) and (28). This immediately reveals the familiar result that the first best is achieved if and only if $i_t = r_t^{n0}$. In the first best, the nominal rate perfectly track the natural real rate. This is true with both rational and behavioral agents.[42,43] It is possible as long as the ZLB does not bind ($r_t^{n0} \geq 0$).

The optimal monetary policy perfectly replicates the flexible-price equilibrium. This is true in the model because monetary policy does not have to substitute for missing tax instruments (Correia et al. (2008)). Hence we recover the traditional, optimistic message of optimal monetary policy in a New Keynesian environment. I defer a discussion of optimal monetary policy with a binding ZLB constraint to Section 5, since a full analysis of policy at the ZLB in the present model requires joint consideration of monetary and fiscal instruments.

4.3 Optimal Policy with Complex Tradeoffs: Reaction to a Cost-Push Shock

The previous shocks (productivity and discount rate shocks) allowed monetary policy to attain the first best. I now consider a shock that doesn't allow the monetary policy to reach the first best, so that trade-offs can be examined. Following the tradition, I consider a "cost-push shock", i.e. a disturbance ν_t to the Phillips curve, which becomes: $\pi_t = \beta M^f \mathbb{E}_t\left[\pi_{t+1}\right] + \kappa x_t + \nu_t$, with ν_t following an AR(1): $\nu_t = \rho_\nu \nu_{t-1} + \varepsilon_t^\nu$.[44,45]

What is the optimal policy then? I examine the optimal policy first if the central bank can commit to actions in the future (the "commitment" policy), and then if it cannot commit (the "discretionary" policy).

---

[42]If the inflation target were $\bar{\pi}$, the nominal rate would be real rate plus inflation target $i_t = r_t^n + \bar{\pi}$. Throughout I assume $\bar{\pi} = 0$ for simplicity.

[43]Sections 4.2-4.3 give optimal policy on the equilibrium path. To ensure determinacy, one simply adds a Taylor rule around it: if the equilibrium path predict values i_t^*, π_t^*, x_t^*, the policy function is: $i_t = i_t^* + \phi_x\left(x_t - x_t^*\right) + \phi_\pi\left(\pi_t - \pi_t^*\right)$ with coefficients ϕ that satisfy the modified Taylor criterion (35).

[44]For instance, if firms' optimal markup increases (perhaps because the elasticity of demand changes), they will want to increase prices and we obtain a positive ν_t (see Clarida et al. (1999) and Galí (2015, Section 5.2)).

[45]Analyzing an early version of the present model, Bounader (2016) examined various constrained policies and derived independently some results in Section 4.3, though not the key result on the non-optimality of price-level targeting.
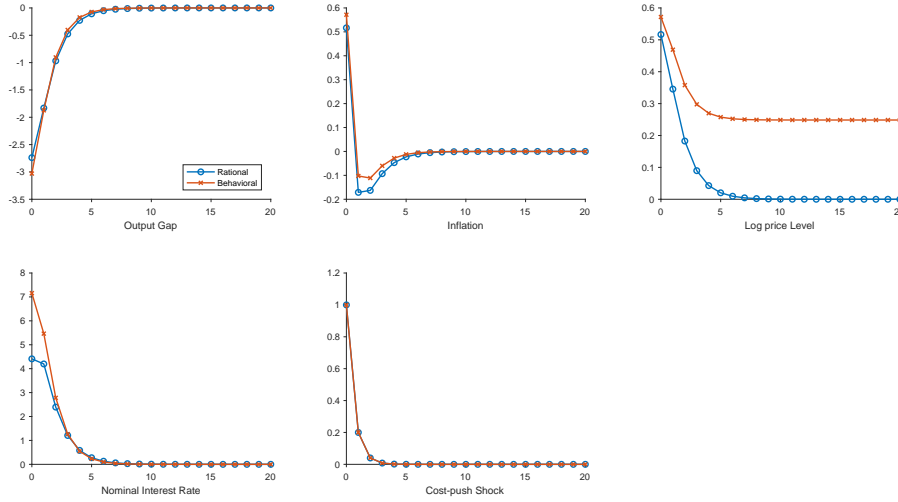
Figure 3: This figure shows the optimal interest rate policy in response to a cost-push shock (ν_t), when the central bank follows the optimal commitment strategy. When firms are rational, the optimal strategy entails "price level targeting", i.e. the central bank will engineer a deflation later to come back to the initial price level. This is not the optimal policy with behavioral firms. This illustrates Proposition 4.2. Units are percentage points. The cost-push shock follows an AR(1) process with autocorrelation $\rho_\nu = 0.2$.

Proposition 4.2 (Optimal policy with commitment: suboptimality of price level targeting) *The optimal commitment policy entails*

$$\pi_t = \frac{-\vartheta}{\kappa} \left(x_t - M^f x_{t-1} \right), \tag{41}$$

so that the (log) price level ($p_t = \sum_{\tau=0}^{t} \pi_\tau$, normalizing the initial log price level to $p_{-1} = 0$) satisfies

$$p_t = \frac{-\vartheta}{\kappa} \left(x_t + \left(1 - M^f \right) \sum_{\tau=0}^{t-1} x_\tau \right). \tag{42}$$

With rational firms ($M^f = 1$), the optimal policy involves "price level targeting": it ensures that the price level mean-reverts to a fixed target ($p_t = \frac{-\vartheta}{\kappa} x_t \to 0$ in the long run).[46] However, with behavioral firms, the price level is higher (even in the long run) after a positive cost-push shock: the optimal policy does not seek to bring the price level back to baseline.

"Price level targeting" and "nominal GDP targeting" are not optimal anymore when firms are behavioral. Price level targeting is optimal with rational firms, but not with behavioral firms. Qualitatively, the commitment to engineer a deflation later helps today, because firms are very forward looking (see Figure 3). That force is dampened in the present behavioral model. The recommendation of price level targeting, one robust prediction of optimal policy model under the rational model, has been met with skepticism in the policy world – in part, perhaps, because its justification isn't very intuitive.[47] This

---

[46]Only M^f, not M, enters in the equilibrium π_t, as the planner maximizes over π_t and x_t welfare (40) subject to $\pi_t = \beta M^f \mathbb{E}_t [\pi_{t+1}] + \kappa x_t + \nu_t$, both of which do not depend on M. The value of M does impact the implementation of the policy, namely i_t.

[47]This is not an intuitive result even in the rational model: in the derivation, this is because the coefficient β in the Phillips curve and the rate of time preference for policy in (40) are the same – something that is not intuitive. That identity is broken in the behavioral model (if the planner instead had a discount factor βM^f, then price level targeting would again be optimal). This is analogous to the Slutsky symmetry in the rational model: there is no great intuition

lack of intuitive justification may be caused by the fact that it is not robust to behavioral deviations, as Proposition 4.2 shows. Likewise, "nominal GDP targeting" is optimal in the traditional model, but is suboptimal with behavioral agents.[48] I next examine the optimal discretionary policy.

Proposition 4.3 (Optimal discretionary policy) *The optimal discretionary policy entails:*

$$\pi_t = \frac{-\vartheta}{\kappa} x_t, \tag{43}$$

so that on the equilibrium path: $i_t = K\nu_t + r_t^n$ *with* $K = \frac{\kappa\sigma^{-1}(1-M\rho_\nu)+\vartheta\rho_\nu}{\kappa^2+\vartheta(1-\beta M^f \rho_\nu)}$.

For persistent shocks ($\rho_\nu > 0$), the optimal policy is less aggressive (K is lower) when firms are more behavioral (when M^f is lower, controlling the value of κ, σ, ϑ).[49] This is because with more myopic firms, future cost-push shocks do not affect much the firms' pricing today, hence the central bank needs to respond less to them.

5   Fiscal Policy

I now study fiscal policy with behavioral agents.

5.1   Cognitive Discounting Generates a Failure of Ricardian Equivalence

For now, fiscal policy means cash transfers from the government to agents and lump-sum taxes (government consumption is zero). Hence, it would be completely ineffective in the traditional model, which features rational, Ricardian consumers. I call B_t the real value of government debt in period t, before period-t taxes. Linearizing, it evolves as $B_{t+1} = R(B_t + \mathcal{T}_t)$, where $\mathcal{T}_t$ is the lump-sum transfer given by the government to the agent (so that $-\mathcal{T}_t$ is a tax).[50] I also define d_t, the budget deficit (after the payment of the interest rate on debt) in period t, $d_t := \mathcal{T}_t + \frac{r}{R}B_t$, so that public debt evolves as:[51]

$$B_{t+1} = B_t + Rd_t. \tag{44}$$

Section 11.1 details the specific assumption I use to capture the agent's worldview. Summarizing, the situation is as follows. Suppose that the government runs a deficit and gives a rebate $\mathcal{T}_t$ to the agents. Agents see the increase in their income, but, because of cognitive discounting, they see only partially the associated future taxes. Hence, they spend some of that transfer, and increase their consumption. The macroeconomic impact of that is as follows.

---

for its justification in the rational model; this is in part because it fails with behavioral agents (Gabaix (2014)). Our intuitions are often (unwittingly) calibrated on our experience as living behavioral agents.

[48]Figure 3 gives some more intuition. Consider the behavior of the interest rate. The interest rate response is milder with rational firms than with behavioral firms. The reason is that monetary policy (especially forward guidance) is more potent with rational firms (they discount the future at β, not at the lower rate $\beta M^f < \beta$), so the central bank can act more mildly to obtain the same effect. In addition, the gains from commitment are lower, as firms don't react much to the future. The optimal policy still features "history dependence" (in the terminology of Woodford (2003b)), even when the cost-push shock has no persistence: see equation (41).

[49]The analogue of Figure 3 for this no-commitment case is in Figure 7 of the Online Appendix.

[50]Without linearization, $B_{t+1} = \frac{1+i_t}{1+\pi_{t+1}}(B_t + \mathcal{T}_t)$, where $\frac{1+i_t}{1+\pi_{t+1}}$ is the realized gross return on bonds. Linearizing, $B_{t+1} = R(B_t + \mathcal{T}_t)$. Formally I consider the case of small debts and deficits, which allows us to neglect the variations of the real rate (i.e. second-order terms $O\left(\left|\frac{1+i_t}{1+\pi_{t+1}} - R\right|(|B_t| + |d_t|)\right)$).

[51]Indeed, $B_{t+1} = R\left(B_t - \frac{r}{R}B_t + d_t\right) = B_t + Rd_t$.

Proposition 5.1 (Discounted Euler equation with sensitivity to budget deficits) *Because agents are not Ricardian, budget deficits temporarily increase economic activity. The IS curve (23) becomes:*

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] + b_d d_t - \sigma\left(i_t - \mathbb{E}_t\left[\pi_{t+1}\right] - r_t^{n0}\right), \tag{45}$$

where r_t^{n0} is the "pure" natural rate with zero deficits (derived in (22)), d_t is the budget deficit and $b_d = \frac{\phi r R(1-\bar{m})}{(\phi+\gamma)(R-\bar{m})}$ is the sensitivity to deficits. When agents are rational, $b_d = 0$, but with behavioral agents, $b_d > 0$. In the sequel, we will write this equation by saying that the behavioral IS curve (24) holds, but with the following modified natural rate, which captures the stimulative action of deficits:

$$r_t^n = r_t^{n0} + \frac{b_d}{\sigma}d_t. \tag{46}$$

Hence, bounded rationality gives both a discounted IS curve and an impact of fiscal policy: $b_d > 0$.[52,53]

Here I assume a representative agent.[54] This analysis complements analyses that assume heterogeneous agents to model non-Ricardian agents, in particular rule-of-thumb agents à la Campbell and Mankiw (1989), Galí et al. (2007), Mankiw (2000), Bilbiie (2008), Mankiw and Weinzierl (2011) and Woodford (2013).[55,56] Models with hand-to-mouth consumers do generate an impact of deficits ($b_d > 0$), but they also yield $M = M^f = 1$.[57] In addition, when dealing with complex situations a representative agent is often simpler. In particular, it allows us to evaluate welfare unambiguously.

5.2 Consequences for Fiscal Policy

Substitutability of monetary and fiscal policy When both monetary and fiscal policy instruments are available, the present model features a generic form of substitutability of the two. In order to see this, I go back to the environment discussed in Section 4.2, in which we have productivity or discount factor shocks (but no cost-push shocks) that alter the natural rate of interest r_t^{n0}. In this case, we can characterize the first-best policy as follows.

---

[52]The Online Appendix (Section 12.8) works out a slight variant, where debt mean-reverts to a fixed constant. The economics is quite similar.

[53]The sensitivity b_d decreases with $\bar{m}$. In the limit $r = \frac{1}{\beta} - 1 \rightarrow 0$ we have that $b_d \rightarrow 0$. This is because the MPC out of income also goes to zero.

[54]There is a lot of debate about Ricardian equivalence. The provisional median opinion is that it only partly holds. For instance, the literature on tax rebates (see Johnson et al. (2006)) appears to support $b_d > 0$.

[55]The model in Galí et al. (2007) is richer and more complex, as it features heterogeneous agents. Omitting the monetary policy terms, instead of $x_t = \mathbb{E}_t\left[\sum_{\tau \geq t} M^{\tau-t} b_d d_\tau\right]$, as obtained by iterating (45) forward, they generate $\hat{c}_t = \Theta_n n_t - \Theta_t t_t^r$. Here n_t is the deviation of employment from its steady state, t_t^r is the log-deviation from steady state of the taxes levied on a fraction of agents who are hand-to-mouth, and Θ_n, Θ_t are positive constants. Hence, one key difference is that in the present model, future deficits matter as well, whereas in their model, they do not.

[56]Mankiw and Weinzierl (2011) have a form of the representative agent with a partial rule of thumb behavior. They derive an instructive optimal policy in a 3-period model with capital (which is different from the standard New Keynesian model), but do not analyze an infinite horizon economy.

[57]Consider the case without fiscal policy. Suppose that a fraction f^h (resp. $f^r = 1 - f^h$) consists of hand-to-mouth (resp. rational) agents who just consume their income $c_t^h = y_t$. Aggregate consumption is $c_t = f^r c_t^r + f^h c_t^h$ and the resource constraint is $y_t = c_t$. But as $c_t^h = y_t$, this implies $y_t = c_t^h = c_t^r$. The hand-to-mouth consume exactly like rational agents. Hence, having hand-to-mouth agents changes *nothing* in the IS equation, and $M = 1$. With fiscal policy, however, those agents do make a difference, i.e. create something akin to $b_d > 0$, but still with $M = 1$.

Lemma 5.2 (First best) *When there are shocks to the natural rate of interest, the first best is achieved if and only if at all dates:*

$$i_t = r_t^n \equiv r_t^{n0} + \frac{b_d}{\sigma} d_t, \tag{47}$$

where r_t^{n0} is the "pure" natural rate of interest given in (22) *and is independent of fiscal and monetary policy. This condition pins down the optimal sum of monetary and fiscal policy (i.e. the value of $i_t - \frac{b_d}{\sigma} d_t$), but not their precise values, as the two policies are perfect substitutes.*

So, if the economy has a lower pure natural interest rate r_t^{n0} (hence "needs loosening"), the government can either decrease interest rates, or increase deficits. Monetary and fiscal policy are perfect substitutes in this model. Their mix is precisely pinned down (i.e. (47) should hold), but the share done via monetary versus (lump-sum) fiscal policy is indeterminate, strictly speaking.[58]

The benchmark of frictionless prices   To clarify ideas, consider the case of fully flexible prices. Suppose that the government changes lump-sum taxes, keeping government consumption to 0. Then, optimality is still determined by a static condition, $N_t^\phi = e^{\zeta_t} c_t^{-\gamma}$ and $c_t = e^{\zeta_t} N_t$, so consumption and labor are independent of fiscal policy. Only the natural rate of interest r_t^n changes. Note that here the assumption of zero capital is key: with capital, a lump-sum tax cut would lead to extra savings, investment, and output. In this sense, a weaker form of Ricardian equivalence holds in this behavioral model without capital: one needs the combination of sticky prices and cognitive discounting in order for deficits to have an output effect.

When the ZLB binds: "Helicopter drops of money" / Fiscal transfers as an optimal cure
When the natural rate becomes negative (and with low inflation), the optimal nominal interest rate is negative, which is by and large not possible. That is the ZLB. The first best is not achievable in the traditional model and the second best policy is quite complex.[59] However, with behavioral agents, there is an easy first best policy:[60]

$$\text{First best at the ZLB: } i_t = 0 \text{ and deficit: } d_t = \frac{-\sigma}{b_d} r_t^{n0}, \tag{48}$$

i.e. fiscal policy runs deficits to stimulate demand.[61] By "fiscal policy" I mean transfers (from the government to the agents) or equivalently "helicopter drops of money", i.e. checks that the central bank might send (this gives some fiscal authority to the central bank).[62] This is again possible because agents are not Ricardian. In conclusion, behavioral considerations considerably change policy at the ZLB, and allow the achievement of the first best.[63]

---

[58]If there are budget deficits, the central bank must "lean against behavioral biases interacting with fiscal policy". For instance, suppose that (for some reason) the government is sending cash transfers to the agents, $d_t > 0$. That creates a boom. Then, the optimal policy is to still enforce zero inflation and output gap by raising interest rates.

[59]The first best is not achievable, and second best policies are complex, as has been analyzed by a large number of authors, e.g. Eggertsson and Woodford (2003), Werning (2012) and Galí (2015, Section 5.4).

[60]A variant is: $i_t = \epsilon$ and $d_t = \frac{-\sigma}{b_d}(r_t^n - \epsilon)$, for some small $\epsilon > 0$, to ensure the determinacy of the Taylor rule on the policy (see footnote 43), which requires the possibility of lowering rates out of the equilibrium path.

[61]Another way of seeing this is that fiscal stimulus raises the natural rate of interest $r_t^n = r_t^{n0} + \frac{b_d}{\sigma} d_t$, and can set it to zero.

[62]The central bank could also rebate the "seigniorage check" to the taxpayers rather than the government, and write bigger checks at the ZLB, and smaller checks outside the ZLB.

[63]In Section 12.9 I analyze a richer situation, and show that the possibility of future fiscal policy can have ex-ante benefits – it makes agents confident about the future, as they know that the government will not run out of tools.

Government spending multiplier greater than one with behavioral agents Finally, let us consider what happens when the government purchases an amount G_t of the aggregate good, and consumes it. I assume that this enters additively in the utility function so that this does not distort the agent's decision. We therefore have $u(c, N, G) = \frac{c^{1-\gamma}-1}{1-\gamma} - \frac{N^{1+\phi}}{1+\phi} + U(G)$ for some function U.

The analytics are fully developed in Section 12.10. One result is easy to state. Suppose that the government purchases at time 0 an amount G_0, financed by a deficit $d_0 = G_0$, and the central bank does not change the nominal rate at time 0. Then the fiscal multiplier is

$$\frac{dY_0}{dG_0} = 1 + b_d, \tag{49}$$

reflecting the fact that government spending has a "direct" effect of increasing GDP one-for-one, and then an "indirect" effect of making people feel richer (as their income increases by G_0 and they don't fully see the increase in future taxes because of cognitive discounting).[64] Hence, this behavioral model generates both $M < 1$ and a fiscal multiplier greater than 1, something that is very hard to generate in models with rational, credit constrained agents (Bilbiie (2018)).

6 Behavioral Enrichments of the Model

In this section I present a number of enrichments of the baseline model laid out so far that are of conceptual and empirical interest. Cognitive discounting captures inattention to the future, but of course people can also be inattentive to contemporaneous variables. Therefore, I first enrich the assumptions of the basic model by allowing for other forms of inattention that induce a "term structure" pattern in the behavioral biases of consumers and producers: I discuss these in Sections 6.1 and 6.2, respectively. These enrichments induce further lessening of the effects of monetary policy as well as a flatter Phillips curve. Second, I present an extension of the model that allows for non-zero or even time-varying trends in inflation. This is empirically realistic and also allows me to address recent paradoxes pointed out by the neo-Fisherian literature.

6.1 Term Structure of Consumer Attention

For conceptual and empirical reasons, I wish to explore the possibility that consumers may perceive certain macro variables more imperfectly than others: for example, consumers may pay overall less attention to the interest rate than they do to income. Theoretically, I found it instructive to see where the intercept, rather than the slope of attention, matters. Also given these various "intercepts of attention" are conceptually natural, they are likely to be empirically relevant as well when future studies measure attention. To capture this, I assume that the agent perceives the law of motion of wealth as:

$$k_{t+1} = \boldsymbol{G}^{k,BR}(c_t, N_t, k_t, \boldsymbol{X}_t) := \left(1 + \bar{r} + \hat{r}^{BR}(\boldsymbol{X}_t)\right)\left(k_t + \bar{y} + \hat{y}^{BR}(N_t, \boldsymbol{X}_t) - c_t\right), \tag{50}$$

---

[64]The decomposition is as follows. The government's action is the sum of (i) a purchase of G_0, financed by a contemporaneous taxes of G_0 and (ii) a tax cut of G_0. The impact of (i) is to increase GDP by G_0 (as is also true in a rational New Keynesian model), and the impact of (ii) is to increase GDP by $b_d G_0$ (as in the present behavioral model). The total effect is to increase GDP by $(1 + b_d) G_0$.

where $\hat{r}^{BR}(\boldsymbol{X}_t)$ and $\hat{y}^{BR}(N_t, \boldsymbol{X}_t)$ are the perceived interest rate and income, given by:

$$\hat{r}^{BR}(\boldsymbol{X}_t) = m_r \hat{r}(\boldsymbol{X}_t), \qquad \hat{y}^{BR}(N_t, \boldsymbol{X}_t) = m_y \hat{y}(\boldsymbol{X}_t) + \omega(\boldsymbol{X}_t)(N_t - N(\boldsymbol{X}_t)), \tag{51}$$

and where m_r, m_y are attention parameters in $[0, 1]$, and $\hat{r}(\boldsymbol{X}_t)$, $\hat{y}(N_t, \boldsymbol{X}_t)$ are the true values of interest rate and personal income, while $\hat{y}(\boldsymbol{X}_t) = \hat{y}(N(\boldsymbol{X}_t), \boldsymbol{X}_t)$ is the true value aggregate income (given aggregate labor supply $N(\boldsymbol{X}_t)$) – all expressed as deviations from the steady state. When m_r, m_y and $\bar{m}$ are equal to 1, the agent is the traditional, rational agent. Here m_r, m_y capture attention to the interest rate and income. For instance, if $m_r = 0$, the agent "doesn't pay attention" to the interest rate – formally, he replaces it by $\bar{r}$ in his perceived law of motion. When $m_r \in [0, 1)$, he partially takes into account the interest rate – really, the deviations of the interest rate from its mean. A microfoundation for the attention parameters m_r and m_y is discussed in Gabaix (2014), Appendix B.[65]

Here $\hat{y}^{BR}(N_t, \boldsymbol{X}_t)$ is his perceived income, and perceived aggregate income is $\hat{y}^{BR}(\boldsymbol{X}_t) = \hat{y}^{BR}(N(\boldsymbol{X}_t), \boldsymbol{X}_t) = m_y \hat{y}(\boldsymbol{X}_t)$: the agent perceives only a fraction of income. However, he correctly perceives that the marginal income is $\frac{\partial}{\partial N} \hat{y}^{BR}(N_t, \boldsymbol{X}_t) = w_t$. This captures that the agent is smart enough to appreciate fully today's marginal impact of working more, though anticipating his total income is harder, especially in the future. Given these perceptions, the agent solves $\max_{(c_t, N_t)_{t \geq 0}} U$ subject to (8) and (50).

Term structure of attention to interest rate and income This formulation, together with Lemma 2.2, implies:[66]

Lemma 6.1 (Term structure of attention) *We have:*

$$\mathbb{E}_t^{BR}\left[\hat{r}^{BR}(\boldsymbol{X}_{t+k})\right] = m_r \bar{m}^k \mathbb{E}_t[\hat{r}(\boldsymbol{X}_{t+k})], \qquad \mathbb{E}_t^{BR}\left[\hat{y}^{BR}(\boldsymbol{X}_{t+k})\right] = m_y \bar{m}^k \mathbb{E}_t[\hat{y}(\boldsymbol{X}_{t+k})]. \tag{52}$$

In words, for the interest rate (the same holds for income):

$$\text{Perceived deviation in } k \text{ periods} = m_r \bar{m}^k \times (\text{True deviation in } k \text{ periods}).$$

Hence, we obtain a "term structure of attention". The factor m_r is the "level" or "intercept" of attention, while the factor $\bar{m}$ is the "slope" of attention as a function of the horizon. The same holds for aggregate income.

Cognitive discounting, which was captured in the main model with the parameter $\bar{m}$, is a form of inattention to future variables. In contrast, the contemporaneous attention parameters m_y and m_r capture limited attention to contemporaneous variables. We see that they typically interact multiplicatively.

If the reader seeks a model with just one free parameter, I recommend setting $m_r = m_y = 1$ (the rational values) and keeping $\bar{m}$ as the main parameter governing inattention, thus recovering the basic formulation of Section 2.

---

[65]Greenwood and Hanson (2015) present a model of industry dynamics in which firms imperfectly perceive the equilibrium responses of their competitors. This competition neglect is a form of inattention to contemporaneous variables similar to the one modeled here and in Section 6.2.

[66]We have:
$$\mathbb{E}_t^{BR}\left[\hat{r}^{BR}(\boldsymbol{X}_{t+k})\right] = \mathbb{E}_t^{BR}[m_r \hat{r}(\boldsymbol{X}_{t+k})] = m_r \mathbb{E}_t^{BR}[\hat{r}(\boldsymbol{X}_{t+k})] = m_r \bar{m}^k \mathbb{E}_t[\hat{r}(\boldsymbol{X}_{t+k})].$$

Consumption and labor supply I now detail the consequences of these enrichments, for a behavioral agent with small initial wealth (Section 11.2 gives the derivation).[67]

Proposition 6.2 (Behavioral consumption function) *In this behavioral model, consumption is:* $c_t = c_t^d + \hat{c}_t$, *with* $c_t^d = \bar{y} + b_k k_t$, $b_k := \frac{\bar{r}}{R}\frac{\phi}{\phi+\gamma}$, *and, up to second order terms:*

$$\hat{c}_t = \mathbb{E}_t\left[\sum_{\tau\geq t}\frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(b_r m_r \hat{r}(\boldsymbol{X}_\tau) + m_Y\frac{\bar{r}}{R}\hat{y}(\boldsymbol{X}_\tau)\right)\right], \tag{53}$$

with $b_r := \frac{-1}{\gamma R^2}$, *and* $m_Y := \frac{\phi m_y+\gamma}{\phi+\gamma}$. *Labor supply satisfies the usual condition* $N_t^\phi = \omega_t c_t^{-\gamma}$, *i.e., in deviations from the steady state,* $\hat{N}_t = \frac{1}{\phi}\hat{\omega}_t - \frac{\gamma}{\phi}\hat{c}_t$. *The policy of the rational agent is a particular case, setting* $\bar{m}, m_r, m_y$ *to 1.*

In (53), consumption reacts to future interest rates and income deviations, dampening future values by a factor $\bar{m}^{\tau-t}$ at horizon $\tau - t$, as in (52). Note that this agent is "globally patient" for steady-state variables. For instance, her marginal propensity to consume wealth is $\frac{\bar{r}}{R}\frac{\phi}{\phi+\gamma}$, like for the rational agent.[68] However, she is myopic to small macroeconomic disturbances in the economy.

The behavioral IS curve with imperfect attention to income and interest rate I next solve for the general equilibrium consequences of policy (53). The resulting IS curve is next (the derivation, in Section 11.2, is instructive, and quite simple).

Proposition 6.3 (Discounted Euler equation, with term structure of attention) *In the enriched model with partial attention to income and interest rate, we obtain a variant of the behavioral IS curve (23) in which* $M = \frac{\bar{m}}{R-rm_Y} \in [0,1]$ *for the macro parameter of attention and* $\sigma := \frac{m_r}{\gamma R(R-rm_Y)} \in \left[0, \frac{1}{\gamma R}\right]$, *with* m_Y *given by* $m_Y = \frac{\phi m_y+\gamma}{\phi+\gamma}$. *In the rational model,* $M = 1$.

Understanding discounting in rational and behavioral models It is now worth pondering where the discounting by M comes from in the behavioral IS curve, via the enriched specification in Proposition 6.3. What is the impact at time 0 of a one-period fall of the real interest rate $\hat{r}_\tau$, in partial and general equilibrium, in both the rational and the behavioral model (as in Angeletos and Lian (2017) and Farhi and Werning (2017))? For simplicity, I take $\hat{r}_\tau^n = 0$ here.

Let us start with the *rational model*. In partial equilibrium (i.e., taking future income as given), a change in the future real interest rate $\hat{r}_\tau$ changes time-0 consumption by the direct (i.e. partial equilibrium) impact (see (53)):

$$\text{Rational agent: } \Delta^{\text{direct}} := \left.\frac{\partial\hat{c}_0}{\partial\hat{r}_\tau}\right|_{(y_t)_{t\geq 0}\text{ held constant}} = -\alpha\frac{1}{R^\tau},$$

where $\alpha := \frac{1}{\gamma R^2}$. Hence, there is discounting by $\frac{1}{R^\tau}$. However, in general equilibrium (i.e., when the

---

[67]I allow for k_t different from 0, as private wealth is non-zero when there is an active fiscal policy.

[68]There is a subtlety here, which Section 12.2 details. The MPC out of wealth is only $\frac{\bar{r}}{R}\frac{\phi}{\phi+\gamma}$, because higher wealth translates into not just more consumption of goods, but also more leisure. However, future booms have an impact of consumption that is $\frac{\bar{r}}{R}$ when $m_y = 1$. This is because they affect the agent's decisions both through higher income, and through higher wages.

impact of $\hat{r}_\tau$ on income flows $(y_t)_{t \geq 0}$ is taken into account), the impact is (see (25) with $M = 1$),

$$\text{Rational agent: } \Delta^{\text{GE}} := \frac{d\hat{c}_0}{d\hat{r}_\tau} = -\alpha R,$$

so that there is no discounting by $\frac{1}{R^{\tau+1}}$. The reason is the following: the rational agent sees the "first round of impact", that is $-\alpha \frac{\hat{r}_\tau}{R^\tau}$; a future interest rate cut will raise consumption. But he also sees how this increase in consumption will increase other agents' future consumptions, hence increase his future income, hence his own consumption: this is the second-round effect. Iterating all other rounds (as in the Keynesian cross), the initial impulse is greatly magnified via this aggregate demand channel: though the first round (direct) impact is $-\alpha \frac{\hat{r}_\tau}{R^\tau}$, the full impact (including indirect channels) is $-\alpha R \hat{r}_\tau$. This means that the total impact is larger than the direct effect by a factor

$$\frac{\Delta^{\text{GE}}}{\Delta^{\text{direct}}} = R^{\tau+1}. \tag{54}$$

At large horizons τ, this is a large multiplier. Note that this large general equilibrium effect relies upon common knowledge of rationality: the agent needs to assume that other agents are fully rational. This is a very strong assumption, typically rejected in most experimental setups (see the literature on the $p-$beauty contest, e.g. Nagel (1995)).

In contrast, in the *behavioral model*, the agent is not fully attentive to future innovations. So first, the direct impact of a change in interest rates is smaller:

$$\text{Behavioral agent: } \Delta^{\text{direct}} := \left. \frac{\partial \hat{c}_0}{\partial \hat{r}_\tau} \right|_{(y_t)_{t \geq 0} \text{ held constant}} = -\alpha m_r \bar{m}^\tau \frac{1}{R^\tau},$$

which comes from (53). Next, the agent is not fully attentive to indirect effects (including general equilibrium) of future polices. This results in the total effect in (25):

$$\text{Behavioral agent: } \Delta^{\text{GE}} := \frac{d\hat{c}_0}{d\hat{r}_\tau} = -\alpha m_r M^\tau \frac{R}{R - rm_Y},$$

so the multiplier for the general equilibrium effect is (as $M = \frac{\bar{m}}{R-rm_Y}$)

$$\frac{\Delta^{\text{GE}}}{\Delta^{\text{direct}}} = \left(\frac{R}{R - rm_Y} \right)^{\tau+1} \in \left[1, R^{\tau+1} \right], \tag{55}$$

and is smaller than the multiplier $R^{\tau+1}$ in economies with common knowledge of rationality. As m_Y becomes smaller, the multiplier weakens: distant changes in interest rates will be very ineffective if agents are extremely myopic. Importantly, when $m_Y = 1$ (i.e., with only cognitive discounting but no attenuation of attention to present income) direct and GE effects are identically dampened, so that we recover the result (54) obtained in the rational model. This is different from the results in Angeletos and Lian (2017) and Farhi and Werning (2017), whose models find an extra dampening of GE compared to direct effects.

6.2 Imperfect Producer Attention to Inflation and Marginal Costs

The behavioral Phillips curve with imperfect attention to inflation and marginal costs
We can similarly consider a symmetric model enrichment on the producer side, allowing firms to have varying degrees of attention to inflation and marginal costs. At time t, the firm perceives the future profit at date $\tau \geq t$ as:

$$v^{BR}\left(q_{it}, \boldsymbol{X}_\tau\right) := v^0\left(q_{it} - m_\pi^f \Pi\left(\boldsymbol{X}_\tau\right), m_x^f \mu\left(\boldsymbol{X}_\tau\right), c\left(\boldsymbol{X}_\tau\right)\right), \tag{56}$$

where v^0 is as in (13). This boundedly rational perceived profit function replaces (14) in the baseline model of Section 2. This means that the firm, when simulating the future, sees only a fraction m_π^f of future inflation $\Pi\left(\boldsymbol{X}_\tau\right)$, and a fraction m_x^f of the future marginal cost $-\mu\left(\boldsymbol{X}_\tau\right)$ (recall that those two quantities have been normalized to have mean 0 at the steady state). When all the m's are equal to 1, we recover the traditional rational firm from the New Keynesian model. The firm's decision problem is now

$$\max_{q_{it}} \mathbb{E}_t^{BR} \sum_{\tau=t}^{\infty} (\beta\theta)^{\tau-t} \frac{c\left(\boldsymbol{X}_\tau\right)^{-\gamma}}{c\left(\boldsymbol{X}_t\right)^{-\gamma}} v^{BR}\left(q_{it}, \boldsymbol{X}_\tau\right). \tag{57}$$

This yields the following optimal price-setting behavior:

$$p_t^* = p_t + (1-\beta\theta) \sum_{k=0}^{\infty} (\beta\theta\bar{m})^k \mathbb{E}_t\left[m_\pi^f\left(\pi_{t+1} + ... + \pi_{t+k}\right) - m_x^f \mu_{t+k}\right], \tag{58}$$

which is analogous to the solution obtained in (26) but now contains the additional scalars m_π^f and m_x^f parameterizing attention to inflation and macro disturbances, respectively, in addition to the overall cognitive discounting factor $\bar{m}$.[69] The general-equilibrium consequences of this model enrichment are as follows:

Proposition 6.4 (Phillips curve with behavioral firms, allowing for imperfect attention to inflation and costs) *In the enriched model with partial attention to inflation and marginal costs, we obtain a variant of the behavioral Phillips curve (28) in which $M^f = \bar{m}\left(\theta + \frac{1-\beta\theta}{1-\beta\theta\bar{m}} m_\pi^f (1-\theta)\right) \in [0,1]$ for the macro parameter governing producer attention and*

$$\kappa = m_x^f \bar{\kappa} \tag{59}$$

for the slope of the Phillips curve, where $\bar{\kappa}$ is the slope in the traditional model with full attention, m_π^f is the firm's attention to contemporaneous inflation, and m_x^f is the firm's attention to contemporaneous macro output conditions. Hence, in the rational model we obtain $M^f = 1$ and $\kappa = \bar{\kappa}$.

The introduction of the new behavioral parameters m_x^f and m_π^f is helpful in several respects. First, a value $m_x^f < 1$ will be helpful in matching the empirical slope κ of the Phillips curve in periods in which it is quite small without resorting to extreme price stickiness.[70,71] Second, going back to the welfare criterion discussion of Section 4, note that holding the measured Phillips curve slope $\kappa = m_x^f \bar{\kappa}$ constant,

---

[69]A proof of equation (58) is in Section 13 in the Online Appendix.

[70]For instance, to generate exactly the same value of κ used in the calibration in Table 1, one could use the parameter configuration $\theta = 0.7$ and $m_x^f = 0.15$.

[71]In a recent paper, Châu (2019) develops a model in which coefficients in the Phillips curve are state-dependent, because of endogenous attention by firms, and find empirical support for the model's predictions.

the relative weight on the output gap ($\vartheta = \frac{\bar{\kappa}}{\varepsilon} = \frac{\kappa}{m_x^f \varepsilon}$) is now higher when firms are more behavioral (when m_x^f is lower). The traditional model gives a very small relative weight ϑ on the output gap when it is calibrated from the Phillips curve – this is often considered a puzzle, which this extension helps alleviate.[72]

6.3 The Model with Non-Constant Trend Inflation

The model so far assumed a constant trend inflation, normalized to 0. Empirical trend inflation present "regimes" that vary from time to time. Here I present an extension of the model that allows for non-zero trend inflation, and, most importantly, allows to think about the Phillips curve and neo-Fisherian paradoxes in that context.

Assumptions and basic model   The analytics will be very simple, but they require a bit of overhead.

Default inflation. When asked to forecast inflation, firms may look at past inflation, as a number of papers have found, in particular Galí and Gertler (1999).[73] To capture this, I call "default inflation", π_t^d, a signal about future inflation that firms form effortlessly. Conceptually, it might be an optimal simple forecast based on past variables, such as $\pi_t^d = \tilde{\mathbb{E}}_t \left[\pi_{t+1} | \left(\pi_\tau, \pi_\tau^{CB} \right)_{\tau \leq t} \right]$, as in Fuster et al. (2012).[74] Here $\tilde{\mathbb{E}}_t$ indicates that this need not be a completely optimal forecast. Here the variable π_t^{CB} is the target inflation announced by the central bank (e.g., 2%).

The specifics of default inflation will not matter much – all we need is *some* default inflation. For concreteness I will use the following functional form:[75]

$$\pi_t^d = (1 - \zeta) \, \bar{\pi}_t + \zeta \bar{\pi}_t^{CB}, \tag{60}$$

where $\bar{\pi}_t$ and $\bar{\pi}_t^{CB}$ are moving averages of past inflation and inflation guidance (i.e., $\bar{\pi}_t = (1 - \eta) \, \bar{\pi}_{t-1} + \eta \pi_{t-1}$ and $\bar{\pi}_t^{CB} = (1 - \eta_{CB}) \, \bar{\pi}_{t-1}^{CB} + \eta_{CB} \pi_{t-1}^{CB}$), though that detailed specification does not matter. What is important is that default inflation puts a weight $\zeta \in [0, 1]$ on the past central bank guidance, π_t^{CB}, and a weight $1 - \zeta$ on past inflation.[76]

Indexation. A number of authors have found that a form of automatic indexation is useful to fit the aggregate data, and not coincidentally, to hit conceptual targets such as long-run Fisher neutrality.[77]

---

[72]Also, going back to the stability criterion (36), note that greater bounded rationality by firms (lower m_x^f) helps achieving determinacy, as the frequency of price changes becomes infinite, $\kappa \to \infty$ (see equation (118)). So to maintain determinacy (and more generally, insensitivity to the very long run), we need both enough bounded rationality and enough price stickiness, in concordance with the finding of Kocherlakota (2016) that we need enough price stickiness to have sensible predictions in long-horizon models.

[73]Galí and Gertler (1999) also present a model with partially backward looking firms, which this section extends. In the notations of this section, their model has $\eta = 1$, $M = 1$, and $\zeta = 0$, which prevents the determinacy analysis below, where $\zeta > 0$ is crucial.

[74]The forecasted variable might be average future inflation, $(1 - \beta \theta) \sum_{k=0}^{\infty} (\beta \theta)^k \pi_{t+k}$.

[75]Form (60) could come from optimal signal extraction – for instance, when inflation has low volatility, ζ will naturally be bigger; but when inflation is variable and the central bank is not trusted, then ζ will be low. That microfoundation could be formalized, but I won't pursue that here.

[76]In support of adaptive expectations, Malmendier and Nagel (2015) estimate empirically an adaptive learning model and find that agents put a weight of 0.6 on experienced inflation in the past. There is not much systematic evidence on the effect of central bank inflation guidance so far. However, Cavallo et al. (2017) find that upon being presented with official statistics about inflation, US consumers put a weight of 0.8 on the new information when forming their posterior.

[77]The traditional New Keynesian model is partially Fisher non-neutral: if inflation is permanently higher, then output is permanently higher: (28) with $M^f = 1$ gives $x = \frac{1-\beta}{\kappa} \pi$. Indexation restores full Fisher neutrality.

I will follow their lead, and assume here full indexation for firms not reoptimizing their prices (like Christiano et al. (2005), and Smets and Wouters (2007)). If a firm does not "actively" adjust its price in a forward-looking manner like in the Calvo model, it just raises it by π_t^d.

Let us now call $\hat{\pi}_t := \pi_t - \pi_t^d$ the deviation of inflation from the default. As the next proposition spells out, then, we are in a world isomorphic to that of the basic model, except we replace π_t by $\hat{\pi}_t$.

Proposition 6.5 (Behavioral New Keynesian model – augmented by a non-zero trend inflation) *In the extended model with non-zero trend inflation, we obtain the following behavioral version of the New Keynesian model. Decompose inflation as:* $\pi_t = \pi_t^d + \hat{\pi}_t$, *where* π_t^d *is default inflation, and* $\hat{\pi}_t$ *is the deviation from default inflation. Then, we have*

$$x_t = M\mathbb{E}_t [x_{t+1}] - \sigma (i_t - \mathbb{E}_t \pi_{t+1} - r_t^n), \tag{61}$$

$$\hat{\pi}_t = \beta M^f \mathbb{E}_t [\hat{\pi}_{t+1}] + \kappa x_t. \tag{62}$$

We recover the same formulation as in the core model (Proposition 2.5); simply, in the Phillips curve (28), we replace inflation π_t *by "deviation from default inflation",* $\hat{\pi}_t$, *which gives* (62).

This encompasses the basic behavioral model of Proposition 2.5, when default inflation is just 0, i.e. $\pi_t^d \equiv 0$, and $\zeta = 1$. Galí and Gertler (1999) microfound backward-looking behavior in the Phillips curve via indexation; Angeletos and Huo (2019) have shown that something related (with $\pi_t^d = \pi_{t-1}$) this can be obtained as a consequence of incomplete information. We will now explore consequences of Proposition 6.5, which also give insight into the implications of these related models.

Consequences of the augmented model   I complete this study with some observations on the behavior of this model augmented with time-varying trend inflation.

Fisher neutrality holds. Fisher neutrality holds in the extended model of Proposition 6.5. Indeed, suppose indeed that long run inflation (and inflation guidance) is $\bar{\pi}$, then the long run nominal rate is $i = r^n + \bar{\pi}$, and the economy is long-run Fisher neutral. Note that this is not the case in the basic model (which assumes that long run inflation is 0, see in (9)), and that the traditional NK model is only partially Fisher neutral (see footnote 77).

Equilibrium determinacy revisited. Is the equilibrium determinate? The next proposition generalizes the earlier criterion (35).[78]

Proposition 6.6 (Equilibrium determinacy with behavioral agents – with backward looking terms) *In the extended model, the equilibrium is determinate only if:*

$$\phi_\pi + \zeta \frac{(1 - \beta M^f)}{\kappa} \phi_x + \zeta \frac{(1 - \beta M^f)(1 - M)}{\kappa \sigma} > 1. \tag{63}$$

Hence, we have a very similar criterion, except for the appearance of ζ, the weight on the central bank guidance.

When monetary policy is passive ($\phi_\pi = \phi_x = 0$), the economy can be determinate in this behavioral model if agents are behavioral enough (low M, low κ perhaps coming from low m_x^f) and if their expectations are anchored enough, e.g. on central bank guidance (high ζ). However, when monetary

---

[78]This proposition states a necessary condition. The necessary and sufficient condition is (63) and a "Routh-Hurwitz auxiliary condition" stated in the Online Appendix (Proposition 13.1). This auxiliary condition is much more minor, and almost automatically valid in practice (see the discussion around Proposition 13.1).

policy is passive traditional models generate non-determinacy, as they violate the criterion (63). This is the case in the traditional New Keynesian model (which has $M = M^f = 1$), in the indexation model of Galí and Gertler (1999) (which has $\zeta = 0$, $M = 1$, $M^f \in [0,1]$, $\pi_t^d = \pi_{t-1}$),[79] and in the typical old Keynesian model (which has $M^f = 0$, $\zeta = 0$).[80]

Speculating somewhat more, this usefulness of "inflation guidance" may explain why central bankers in recent years did not wish to deviate from an inflation target of 2% (and go to a higher target, say 4%, which would leave more room to avoid the ZLB). They fear that "inflation expectations will become unanchored", i.e. that ζ will be lower: agents will believe the central bank less (as it "broke its word"), which in turn can make the economy equilibrium indeterminate, by (63). This reasoning relies on agents' bounded rationality.

The 1970s The stagflation of the 1970s has been attributed by Clarida et al. (2000) to a violation of the Taylor criterion – in essence, $\phi_\pi < 1$. But we have seen that Japan has arguably $\phi_\pi = 0$, and that this can still be consistent with a determinate equilibrium. How to reconcile these *prima facie* contradictory facts? In the present model, the 1970s can be interpreted as a moment where agents did not believe the central bank enough, i.e. ζ was too low (in part because inflation was volatile, central bank credibility was eroded) – while in Japan, ζ is high enough. Together with the failure of the Taylor criterion documented in Clarida et al. (2000), this leads to criterion (63) being violated. This again suggests that studying empirically the parameter ζ (i.e. the weight of long run inflation guidance by the central bank), and its deeper information-theoretic foundation, would be quite interesting.

Neo-Fisherian experiment: A permanent shock to target inflation I now use the model to analyze neo-Fisherian effects with behavioral agents. I assume that the central bank announces at time 0 an immediate, permanent, unexpected rise of 1% in the nominal rate and in its corresponding target inflation ($i_t = 1\%$ at all dates $t \geq 0$, and the central bank guidance is the corresponding long term target, $\pi_t^{CB} = 1\%$).[81] Figure 4 shows the result.[82] On impact, there is a recession: output and inflation are below trend. However, over time default inflation increases: as the central bank gives "guidance", inflation expectations are raised. In the long run, for this calibration, we obtain Fisher sign neutrality.

This effect is hard to obtain in a conventional New Keynesian model.[83] Cochrane (2017, p.3) summarizes the situation:[84]

> "The natural starting place in this quest [for a negative short-run impact of interest rates on inflation] is the simple frictionless Fisherian model, $i_t = r + \mathbb{E}_t \pi_{t+1}$. A rise in interest rates i produces an immediate and permanent rise in expected inflation. In the search for

---

[79]With those values, the Phillips curve (61) is actually the limit of their Phillips curve when their β is 1.

[80]For instance, the old Keynesian model features a deflationary spiral, because it has $\zeta = 0$. However, we see that in the old Keynesian model, augmented with $\zeta > 0$ (i.e., agents listen enough to the central bank when forming expectations), we can verify the criterion.

[81]To ensure determinacy, we can just add a Taylor rule around that the equilibrium path reported here, as in footnote 43.

[82]In addition to the basic calibration of Table 1, I use also parameters for default inflation: $\eta = 0.5$, $\eta_{CB} = 0.05$, $\zeta = 0.8$.

[83]For instance, in the traditional New Keynesian model a permanent change in the inflation target (i.e., of the intercept j_t of the Taylor rule) involves no transitional dynamics: it leads to an instantaneous jump of the whole economy to the new steady state, and this rise in interest rates leads to a one-for-one rise in inflation. Slow transition dynamics emerge only when departing from the basic model, e.g. by assuming imperfect information (Erceg and Levin (2003)) or sticky wages and/or indexation (Ascari and Ropele (2012)).

[84]This can depend on which equilibrium is selected, leading to some cacophony in the dialogue.

Figure 4: Impact of a permanent rise in the nominal interest rate. At time 0, the nominal interest rate is permanently increased by 1%. The Figure traces the impact on inflation and output. Units are percentage points.
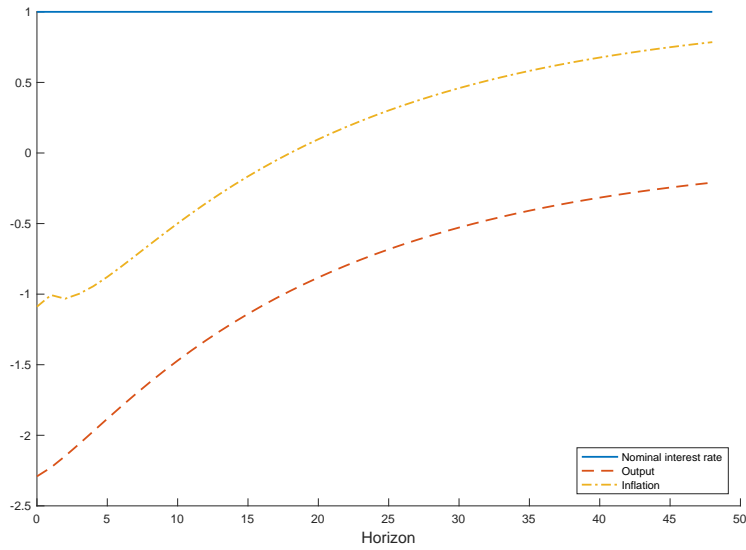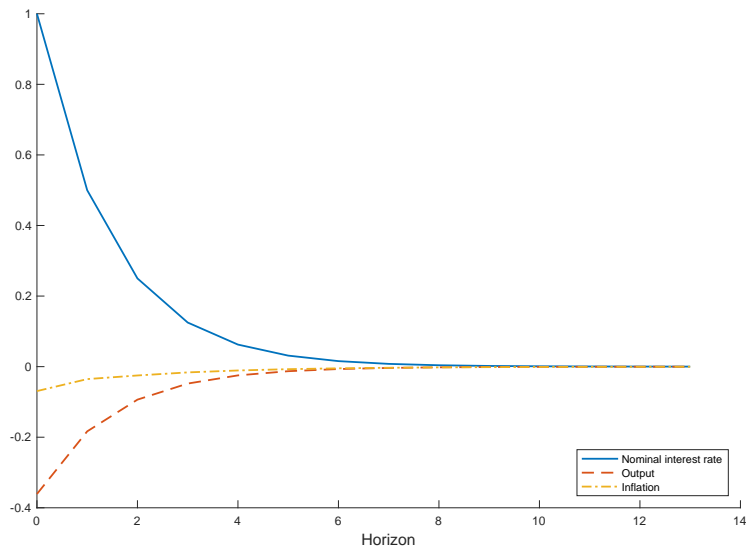

Figure 5: Impact of a temporary rise in the nominal interest rate. At time 0, the nominal interest rate is temporarily increased by 1%. The Figure traces the impact on inflation and output. Units are percentage points.

> a temporary negative sign [one can add] to this basic frictionless model: 1) new Keynesian pricing frictions, 2) backwards-looking Phillips curves, 3) monetary frictions. These ingredients robustly fail to produce the short-run negative sign."

This paper gives a way to overturn this result, coming from agents' bounded rationality. In this behavioral model, raising rates permanently first depresses output and inflation, then in the long run raises inflation (as Fisher neutrality approximately holds), via the credible inflation guidance.

This analysis, of course, is not an endorsement, as this policy is not first best, and leads to a prolonged recession.

A temporary shock to the interest rate. I now study a temporary increase of the nominal interest rate, $i_t = \rho_i^t i_0$ for $t \geq 0$, following an immediate, unexpected rise of 1% in the nominal interest rate at time 0. As the long run is not modified, I assume an inflation guidance of 0, $\pi_t^{CB} = 0$, and use $\rho_i = 0.5$. Figure 5 shows the result. On impact, inflation and output fall, and then mean-revert. The behavior is very close to what happens in the basic model of Section 2.

7 Discussion of the Behavioral Assumptions

Here I discuss at greater length the behavioral assumptions of the model, especially the key Assumption 2.1.

Microeconomic evidence There is mounting microeconomic evidence for the existence of inattention to small dimensions of reality (Gabaix and Laibson (2006), Brown et al. (2010), Caplin et al. (2011), Gabaix (2019)) including taxes (Chetty et al. (2009), Taubinsky and Rees-Jones (2017)), and macroeconomic variables (Coibion and Gorodnichenko (2015)). It is represented in a compact way by the inattention parameters – that is, the m's.[85] This paper also highlights another potential effect that has not specifically been investigated: a "slope" of inattention captured by $\bar{m}$, whereby agents perceive more dimly things that are further in the future. There is no direct evidence about cognitive discounting per se, and one interesting research question would be to investigate this more directly, particularly the "term structure of attention" of Lemma 6.1. The Online Appendix (Section 12.5) discusses extant evidence. It finds that cognitive discounting can explain some of the evidence (aggregate forecast revisions predict aggregate forecast errors on average), in particular from Coibion and Gorodnichenko (2015), if viewed as a theory of the aggregate forecaster. However, the theory would need to be supplemented by other modeling features to match other aspects of the evidence. For example, to match the slow adjustment dynamics of forecast errors to shocks studied in Coibion and Gorodnichenko (2012), one can explicitly model delayed reaction to news (Gabaix and Laibson (2002); Mankiw and Reis (2002)). In order to explain potential over-reaction at the individual forecaster level (as in Bordalo et al. (2018)) one could incorporate overconfidence in the precision of signals. Gabaix and Laibson (2017) argue that a large fraction of the literature on hyperbolic discounting reflects a closely related form of cognitive discounting. One can hope that future research will investigate all this empirically, especially for actual consumers: most data comes from professional forecasters, who are likely to be more rational than the consumers in this model.

---

[85]In this paper, the theme is that of underreaction. It is possible to generate overreaction: if people overestimate the autocorrelation of productivity or income shocks (because it's higher in their default model), they will overreact to them. See Gabaix (2019), Section 2.3.13.

Theoretical microfoundations Section 9 discusses in detail the microfoundations of the inattention parameters, and proposes an endogenization for them. Here I give a summary of the situation. First, pragmatically, my preferred interpretation is that the formulation (8) can be taken as a useful idealization of the agent's simulation process. This is in line with much behavioral economics, in which a plausible description of the thought process is posited, and its consequences analyzed – but the research on its deeper microfoundations is left for the future (for instance, loss aversion is observed and modeled, but there is still no agreement about its "deep" microfoundations, so that loss aversion is directly used, rather than its more remotely speculative microfoundations; and likewise for hyperbolic discounting, fairness etc.). Adjusting for the different stakes, this is similar for, say, equilibrium. One starts with a notion of equilibrium (supply equals demand, or Nash equilibrium), but the hypothetical nanofoundations for how the market will reach equilibrium (e.g. tâtonnement) are typically done in separate studies, and not actively used when thinking about the consequences of equilibrium for concrete economic analyses.

Section 9.1 proposes such a potential nanofoundation: formulation (8) (and the extension (50)) can be viewed as the "representative agent" version of a model in which the agent performs a mental simulation of the future, but receives only noisy signals about that simulation.[86] Importantly, one can allow agents to optimize their attention (Section 9.2). Then, the optimal level of attention reacts to the incentives to pay attention that the agent faces.

Lucas critique In most of this paper the attention parameters are taken to be constant. But for completeness Section 9.2 discusses their endogenization. Attention will not change if, for instance, the volatility of the environment increases by a small or moderate amount (the "sparsity" feature of the theoretical microfoundation is useful for that, as it makes the agent locally non-reactive to things), but it will rise if the volatility of the environment increases a lot (see in particular the discussion after Proposition 9.1). Hence, the Lucas critique does not apply for small or moderate changes, but it does apply to large changes.

Long run learning Relatedly, the agent has forever a biased model of the world (biased by cognitive discounting) – in that sense, she does not learn in the long run. This makes sense, as attention is costly. We do sail through life without learning many things – for example, most people lead happy lives without learning quantum mechanics. Quantum mechanics is difficult, and not crucial to leading a good life. Likewise, in this model, fully understanding interest rates is difficult, and not crucial for life. Learning and attention are effortful, and typically we do not learn all things within a human lifespan.

New degrees of freedom This model is quite parsimonious: there is just one non-standard parameter, the cognitive discounting parameter $\bar{m}$. Section 6 presents enrichments that introduce other behavioral parameters, but these are much less important, and can be disciplined via measurement.[87] In other contexts such as tax salience (Taubinsky and Rees-Jones (2017)), attention parameters are

---

[86]For the welfare part of this paper, it is expedient to have a model in which the representative agent construct holds exactly (it would be interesting to study welfare when agents differ because of the different noisy signals that they receive, but I leave it to future research). Under the first interpretation, this is automatic. Under the second interpretation, we can posit that the agent is really a continuous family of such agents, each of whom takes an infinitesimal decision on consumption and labor supply, so that the representative agent perspective holds.

[87]There is "meta" degree of freedom – where do the m's go? I note that this "meta" problem is present in all of economics. For example, in information economics, it's normally assumed that the agent knows almost all the world perfectly, and has imperfect information about just one or a few variables. Likewise, we introduce adjustment costs in

being progressively better measured (see Gabaix (2019) for a survey), and one can hope that the same thing will happen for macro parameters of attention.

Reasonable variants   Like any model, the framework admits a large number of reasonable variants. I have explored a number of them, and the economics I present here reflects what is robust in those variants.[88] The model here presents one such set of assumptions – essentially, I chose them by aiming at a happy balance between tractability, parsimony and psychological and macroeconomic realism.

8   Conclusion

This paper gives a simple way to think about the impact of bounded rationality on monetary and fiscal policy. Furthermore, we have seen that there is empirical support for the main non-standard elements of the model. As shown in the prior literature, the empirical Phillips curve is partially myopic, so is the IS curve, and agents are partially non-Ricardian. This paper leads to a large number of natural questions.

Theory. I have studied only the most basic model. Doing a similar exploration for its richer variants would be very interesting and relevant both empirically and conceptually: examples include capital accumulation, a more frictional labor market, distortionary taxes, and agents that are heterogeneous in wealth or rationality. The tractable approach laid out in this paper makes the exploration of those questions quite accessible. Relatedly, it facilitates studying optimal central bank policy with behavioral agents under varied situations.[89] An obvious practically useful project would be to enrich currently-used model such as Smets and Wouters (2007) with cognitive discounting and related limited attetion parameters, and estimate such a model.

Empirics. The present work suggests a host of questions for empirical work. One would like to estimate the intercept and slope of attention (i.e. attention to current variables, and how the understanding of future variables decreases with the horizon) using individual-level dynamics for consumers (equation (53)), for firms (equation (58)), and of the whole equilibrium economy (Proposition 2.5). One side-payoff of this work is to provide a parametrized model where these forces can be empirically assessed (by measuring the various m's in the economy).[90]

Surveys. This work also suggests new questions for survey design. One would like to measure people's subjective model of the world – which, like that of this model's agents, may not be accurate. For instance, one could design surveys about people's understanding of impulse-responses in the economy. They would ask questions such as: "Suppose that the central bank raises the interest rate now [or in a year, etc.], what do you think will happen in the economy? How will you change your consumption today?". In contrast, most work assesses people's predictions of individual variables (e.g. Greenwood

---

just one of a few variables, not to all. The modeler chooses which those are – guided by a sense of "what is relevant and interesting". I try to do the same here.

[88]For instance, the agent might extrapolate too much from present income: this gives a high MPC out of current income, but otherwise the macro behavior does not change much (see Section 12.3). She might also suffer from nominal illusion in her perception of the interest rate (Section 12.6). Also, if we had growth, the agent would cognitively discount the deviations $\boldsymbol{X}_t$ from the balanced growth path (see Section 12.7). One could also imagine a number of variants, e.g. (8) and (10) might replace $\bar{m}$ by a diagonal matrix $diag\,(\bar{m}_i)$ of component-specific cognitive discounting factors.

[89]See Nakata et al. (2019) and Benchimol and Bounader (2018).

[90]See Coibion and Gorodnichenko (2015), Afrouzi (2017) and Fuhrer (2017) for progress on those issues.

and Shleifer (2014)) rather than their whole causal model.[91] The parameterization in the present work allows for a way to explore potentially important deviations of the model from the rational benchmark, and suggests particular research designs that focus on the key differential predictions of a rational versus a behavioral model.[92]

In conclusion, this paper offers a parsimonious way to think through the impact of bounded rationality on monetary and fiscal policy, both positively and normatively. It suggests a number of theoretical and empirical questions that would be fruitfully explored.

9 Appendix: Microfoundations for Cognitive Discounting

There are three questions when handling a behavioral model of the type presented here.

1. How does the model generate in a coherent way the agent's consumption and labor supply policies (given attention m)? How does this affect economic outcomes?

2. Is there a story for why we would achieve that formulation?

3. How does the parameter m vary with incentives?

In my view, question 1 is the most important "practical" question – as it is crucial to handle demand functions. Accordingly, I detail it throughout the paper.

Question 2 is addressed in Section 9.1. For practical purposes, it is probably the least crucial. One perhaps useful example is the concept of "equilibrium prices". The way over 99.9% of economics proceeds is to "assume that the market clears at price p", and solve for the price. There is a small and worthy part of economics that thinks about microfoundations for the possibility of finding the equilibrium price (e.g. tâtonnement) – that is a microfoundation that is useful to know, but it is not necessary to repeat it in each paper. On top of that, it is not clear that we have found the true microfoundations for market equilibrium.[93] Rather, those generate only something close, but not exactly equal to, the costless, instantaneous jump to market equilibrium. Hence, practicing economists are aware of those candidate microfoundations for "equilibrium prices", but seldom actively use them. Still, it is healthy for concrete economics to have it as a benchmark in the background. It is in that spirit that I develop a microfoundation in 9.1.

Question 3 is useful in some applications, and I detail it in Section 9.2. The upshot is this: in the "sparsity" framework, there is local rigidity and global flexibility. *Local rigidity*: when variances of primitive shocks remain within a certain bound,[94] then attention parameters remain exactly constant (there, "sparsity" is particularly useful). *Global flexibility*: However, if variances become very high, then attention does increase. The Lucas critique applies to large changes in the environment, but not to small ones. Hence, this paper will work with constant attention – that corresponds to variances that can vary moderately.

---

[91]E.g. it asks questions like: "Are you optimistic about the economy today?" or "Where do you think the economy will be in a year?". See Carvalho and Nechio (2014) for people's qualitative understanding of policy.

[92]E.g. one could ask "Suppose the central bank lowers interest rates by 1% [or the government gives $1000 to all agents] for one period in eight quarters, what will happen to the rest of the economy, and to your decisions?", plot the impulse response, vary the parameter "eight", and compare that to the rational and behavioral models.

[93]The same holds for related concepts, e.g. "Nash equilibrium" and "rational expectations equilibrium". There are "infinitely iterated rounds of reasoning" stories to microfound those, and they generate something like those equilibrium concepts only under very strong and idealized conditions.

[94]For instance, $\mathbb{E}\left[\zeta_t^2 + j_t^2\right] \leq K$, for some bound K. Recall that j_t denotes innovations to monetary policy.

9.1 A Possible "Noisy Simulations" Foundation for Cognitive Discounting

Here is a possible "noisy simulations" microfoundation for cognitive discounting (8), i.e. the fact that the agent perceives:[95]

$$\boldsymbol{X}_{t+1} = \bar{m}\boldsymbol{G^X}\left(\boldsymbol{X}_t, \boldsymbol{\epsilon}_{t+1}\right). \tag{64}$$

One period: basic idea To clarify ideas, let us consider a simpler scenario where at time $t = 0$ the agent simply simulates $\boldsymbol{X}_1$. The true next-period value is $\boldsymbol{X}_1 = \boldsymbol{G^X}\left(\boldsymbol{X}_0, \boldsymbol{\epsilon}_1\right)$. However, the agent receives a noisy signal $\boldsymbol{Y}_1$ about this:

$$\boldsymbol{Y}_1 = \begin{cases} \boldsymbol{X}_1 & \text{with probability } q, \\ \boldsymbol{X}_1' & \text{with probability } 1 - q. \end{cases} \tag{65}$$

That is, with a probability q, he receives the correct value $\boldsymbol{X}_1$, while with probability $1 - q$, there is a "random reset" in the agent's simulation process: then the agent receives a random i.i.d. draw $\boldsymbol{X}_1'$ from the unconditional distribution of $\boldsymbol{X}_1$.[96] This reset captures a form of disruption in the reasoning process.

Normalizing throughout the unconditional mean of $\boldsymbol{X}_1$ to be $\bar{\boldsymbol{X}} = 0$, this implies that the conditional mean given the signal $\boldsymbol{Y}_1$ is:[97]

$$\boldsymbol{X}_1^e\left(\boldsymbol{y}_1\right) := \mathbb{E}\left[\boldsymbol{X}_1 | \boldsymbol{Y}_1 = \boldsymbol{y}_1\right] = q\boldsymbol{y}_1, \tag{66}$$

and given

$$\mathbb{E}\left[\boldsymbol{Y}_1 | \boldsymbol{X}_1\right] = q\boldsymbol{X}_1 + (1 - q)\,\mathbb{E}\left[\boldsymbol{X}_1'\right] = q\boldsymbol{X}_1 + (1 - q)\,\bar{\boldsymbol{X}} = q\boldsymbol{X}_1$$

the average perceived value given the truth is:

$$\bar{\boldsymbol{X}}_1^e\left(\boldsymbol{X}_1\right) := \mathbb{E}\left[\boldsymbol{X}_1^e\left(\boldsymbol{Y}_1\right) | \boldsymbol{X}_1\right] = \mathbb{E}\left[q\boldsymbol{Y}_1 | \boldsymbol{X}_1\right] = q\mathbb{E}\left[\boldsymbol{Y}_1 | \boldsymbol{X}_1\right] = q^2\boldsymbol{X}_1.$$

Hence, defining

$$\bar{m} := q^2 \tag{67}$$

we have

$$\bar{\boldsymbol{X}}_1^e\left(\boldsymbol{X}_1\right) = \bar{m}\boldsymbol{X}_1 = \bar{m}\boldsymbol{G^X}\left(\boldsymbol{X}_0, \boldsymbol{\epsilon}_1\right). \tag{68}$$

This way, the agent will perceive $\bar{m}\boldsymbol{X}_1$ on average. Then, the representative agent (who averages over all agents) will behave according to (64).

---

[95]Here I give a microfoundation for cognitive discounting, which involves iterated simulations of the future. There is also a more static "noisy perceptions" microfoundation for the intercept parameters m_r, m_y, m_x^f, m_π^f: it is detailed in Gabaix (2014), Appendix B.

[96]One could use other formulations, e.g. $\boldsymbol{Y}_1 = q\boldsymbol{X}_1 + \sqrt{1 - q^2}\boldsymbol{X}_1'$. It works the same way (i.e., $\mathbb{E}\left[\boldsymbol{X}_1 | \boldsymbol{Y}_1\right] = q\boldsymbol{Y}_1$, so $\bar{m} = q^2$), but then $\boldsymbol{X}_1$ needs to be Gaussian distributed.

[97]Indeed, calling $g\left(\boldsymbol{X}_1\right)$ the distribution of $\boldsymbol{X}_1$, the joint density of $\left(\boldsymbol{X}_1, \boldsymbol{Y}_1\right)$ is $f\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right) = qg\left(\boldsymbol{y}_1\right)\delta_{\boldsymbol{y}_1}\left(\boldsymbol{x}_1\right) + (1 - q)\,g\left(\boldsymbol{y}_1\right)g\left(\boldsymbol{x}_1\right)$, where δ is the Dirac function. So, as $\int \boldsymbol{x}_1 g\left(\boldsymbol{x}_1\right)d\boldsymbol{x}_1 = 0$,

$$\mathbb{E}\left[\boldsymbol{X}_1 | \boldsymbol{Y}_1 = \boldsymbol{y}_1\right] = \frac{\int \boldsymbol{x}_1 f\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right)d\boldsymbol{x}_1}{\int f\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right)d\boldsymbol{x}_1} = \frac{qg\left(\boldsymbol{y}_1\right)\boldsymbol{y}_1 + (1 - q)\,g\left(\boldsymbol{y}_1\right)\int \boldsymbol{x}_1 g\left(\boldsymbol{x}_1\right)d\boldsymbol{x}_1}{g\left(\boldsymbol{y}_1\right)} = q\boldsymbol{y}_1.$$

How to get all agents to do exactly like the representative agent? For welfare, it is useful for exact representative-agent aggregation to hold. I show two ways to do that: via the family metaphor, and via an "integration within the mind" metaphor. The first yields the linearized version of cognitive discounting, (9), while the second story yields the full non-linear version of cognitive discounting, (8). In laying out foundations, it may be useful to have two potential ways to think about things, so I present both.

The "family metaphor". Suppose each agent is really a "family", made of a continuum of agents $j \in [0,1]$. Agent j takes the optimal action[98] (expressed as a deviation $\hat{a}$ from the steady state optimal action, $\bar{a}$) given his perception, so $\hat{a}_j = a_{\boldsymbol{X}} \mathbb{E}[\boldsymbol{X}_1|\boldsymbol{Y}_{1j}] = a_{\boldsymbol{X}} q \boldsymbol{Y}_{1j}$. The total action of the aggregate agent is then $\hat{a} = \int_0^1 \hat{a}_j dj = a_{\boldsymbol{X}} q \int_0^1 \boldsymbol{Y}_{1j} dj = a_{\boldsymbol{X}} q^2 \boldsymbol{X}_1$:

$$\hat{a}(\boldsymbol{X}_1) = \bar{m} a_{\boldsymbol{X}} \boldsymbol{X}_1. \tag{69}$$

This way, the aggregate agent is the representative agent, at least to the first order. His action is $a(\boldsymbol{X}_1) = \bar{a} + \bar{m} a_{\boldsymbol{X}} \boldsymbol{X}_1$.

The "integration within the mind" metaphor. The agent runs a continuum of simulations $j \in [0,1]$, i.e. obtains draws $\boldsymbol{Y}_{1j} = \boldsymbol{Y}_1(\boldsymbol{X}_1, s_{1j})$, where s_{1j} indexes the simulation j. Each simulation j leads to a posterior mean $\boldsymbol{X}_1^e(\boldsymbol{Y}_{1j}) = q \boldsymbol{Y}_{1j}$. Then, the mind uses an average of those posteriors:

$$\boldsymbol{X}_1^{BR} = \int_0^1 \boldsymbol{X}_1^e(\boldsymbol{Y}_{1j}) \, dj = q^2 \boldsymbol{X}_1 = \bar{m} \boldsymbol{X}_1 = \bar{m} \boldsymbol{G}^{\boldsymbol{X}}(\boldsymbol{X}_0, \boldsymbol{\epsilon}_1).$$

This way, the perceived law of motion is exactly (64).

This law of motion is perceived for a given $\boldsymbol{\epsilon}_1$. If the agent just cares about the mean value of a linearized system, she does this for one value, $\boldsymbol{\epsilon}_1 = 0$. If the agent cares about the whole non-linear system, the procedure is done for all $\boldsymbol{\epsilon}_1$.[99,100]

This completes the microfoundation for a case where the agent simulates the next period.

Several periods Now that the one-period simulation is in place, it is easy to generalize to several periods. We have seen how a value $\boldsymbol{X}_0$ leads to a value of $\boldsymbol{X}_1$ which follows (8). Now, the agent does the same at all periods. She does this going from $\boldsymbol{X}_1$ to $\boldsymbol{X}_2$, etc. By induction, the agent perceives (8) for all dates t.

9.2 Endogenizing Attention

General formulation The traditional New Keynesian model takes pricing frictions as given, and then studies their consequences. One can also endogenize the size of the pricing friction (of θ, see Kiley (2000)), but most of the analysis is most cleanly done by taking the pricing friction as given. Likewise,

---

[98]This action can be any action, e.g. consumption or labor supply.

[99]Implicitly, we assume that the mind can only integrate by taking the sample mean of the signal. It could conceivably take a more sophisticated procedure. When we model bounded rationality (as opposed to optimal information processing), there is a point at which the sophistication of the algorithm must stop. For instance, take level-k models: suppose a level-k agent with $k = 1$. Then, given her signal for the reaction at $k = 1$, she could optimize some more and find some better estimate of the optimal action. Level k models assume that the agent just stops there.

[100]I do not claim that agents do this exactly. I just delineate a stylized scenario that would generate (64). Understanding exactly how people calculate expected values (e.g. approximate $\int V(\varepsilon) f(\varepsilon) d\varepsilon$ for some value V and distribution f) would be very interesting, but completely outside the scope of this study.

in this paper I take the degree of inattention as given, and study its consequences. In this section I sketch how to endogenize it, drawing heavily on Gabaix (2014) and Gabaix (2016).

I show how to endogenize a parameter I call m – this could be $\bar{m}$, or the attention to the interest rate, or to inflation (m_r, m_π^f), and so forth. Call a_t the action at time t, and $\boldsymbol{S}_t$ the state vector. For instance, in the consumer's problem, $a_t = (c_t, N_t)$ (consumption and labor supply) and $\boldsymbol{S}_t = (k_t, \boldsymbol{X}_t)$ (k_t is the agent's personal wealth, which will be 0 in equilibrium in the model without taxes, and $\boldsymbol{X}_t$ is vector of macro variables). The value at the default state is normalized to be $\boldsymbol{S}^d = 0$.

The agent has a subjective value function $V(\boldsymbol{S}, m)$ that is the traditional, rational value function under the subjective model parameterized by m. So, at time t, the agent wishes to maximize

$$v(a_t, \boldsymbol{S}_t, m) = u(a) + \beta \mathbb{E} V\left(\boldsymbol{G}^{\boldsymbol{S}}\left(\boldsymbol{S}_t, a_t, m, \varepsilon_{t+1}\right), m\right), \tag{70}$$

where $V(\boldsymbol{S}, m)$ is the subjective value function, i.e. the value function corresponding to the agent's subjective model of the world – parameterized by m, i.e. the one with the transition function $\boldsymbol{G}(\boldsymbol{S}_t, a_t, m, \varepsilon_{t+1})$. She takes the action

$$a(m, \boldsymbol{S}_t) = \underset{a}{\operatorname{argmax}}\, v(a, \boldsymbol{S}_t, m). \tag{71}$$

Conceptually, the agent would like to maximize true utility, given the imperfect action $a(m, \boldsymbol{S})$, net of costs $\mathcal{K}g(m - m^d)$:

$$\max_m \mathbb{E} v\left(a(m, \boldsymbol{S}_t), \boldsymbol{S}_t, 1\right) - \mathcal{K}g\left(m - m^d\right), \tag{72}$$

i.e. the inclusive utility is evaluated under the true model (indexed by $m = 1$), and the agent wants to avoid paying the thinking costs $\mathcal{K}g(m - m^d)$, with $\mathcal{K} \geq 0$ (with $\mathcal{K} = 0$ being the rational case). Here m^d is a "default" attention, processed for free by the agent. Typically in this paper, $m^d > 0$: we have "for free" some understanding of the future. The thinking cost $g(\cdot)$ can be traced back in turn to the more primitive simulation technology used by the agent in Section 9.1. A typical functional form is the parametrization $g(m - m^d) = |m - m^d|$, which gives an attention function that is always continuous, and constant in part of its domain (Gabaix (2014)).

The problem in (72) is typically intractable (both for the researcher and, presumably, for the agent), so some alternative formulation is needed. The sparse max model in Gabaix (2014) proposes a way to address that difficulty. There, the agent solves a linear-quadratic approximation of this problem (72), taking a Taylor expansion of the utility losses when evaluating optimal attention – but keeping the true nonlinear utility when taking her action, as in (71).[101] This means that (72) is replaced by:[102]

$$\max_m -\frac{1}{2}\Lambda(1 - m)^2 - \mathcal{K}g\left(m - m^d\right), \tag{73}$$

with $\Lambda = \lambda \sigma_{\boldsymbol{S}}^2$,

$$\lambda = -\mathbb{E}\left[\boldsymbol{S}_t^\circ{}' a_{m,\boldsymbol{S}}'\left(m^d, 0\right) v_{aa}\left(a\left(m^d, 0\right), 0, m^d\right) a_{m,\boldsymbol{S}}\left(m^d, 0\right)\boldsymbol{S}_t^\circ\right], \tag{74}$$

where $a_{m,\boldsymbol{S}}$ is the second-order cross partial derivative, and I scale $\boldsymbol{S}_t - \boldsymbol{S}_t^d = \sigma_{\boldsymbol{S}}\boldsymbol{S}_t^\circ$: a higher $\sigma_{\boldsymbol{S}}$ parameterizes the volatility of $\boldsymbol{S}_t$. Here, $\frac{1}{2}\Lambda(1 - m)^2$ is the leading term in the Taylor expansion of

---

[101]The assumptions on $\boldsymbol{G}^{\boldsymbol{S}}(\boldsymbol{S}, a, m)$ imply that $a(m, 0)$ is independent of m: attention affects only the deviations from the default state.

[102]This is generalizes to Gabaix (2014), Definition 1 and Lemma 2. The derivation is in Section 12.11.
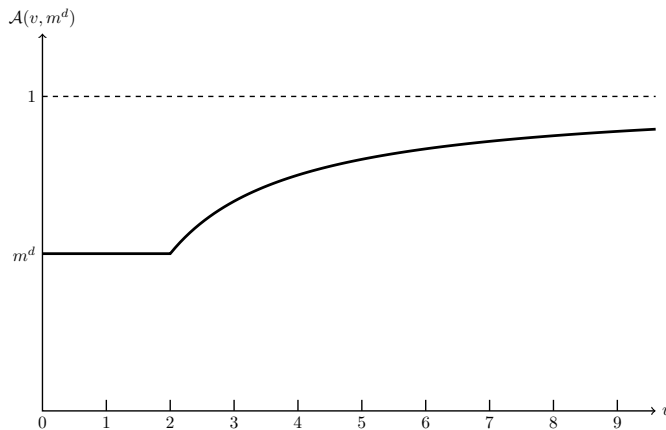
Figure 6: This Figure plots the attention function (75). When the volatility v of the environment is small or moderate, attention is at its default value m^d. However, when volatility increases a lot, attention increases, asymptotically toward 1.

utility losses. The solution is $m = \mathcal{A}\left(\frac{\lambda\sigma_S^2}{\mathcal{K}}, m^d\right)$ with:

$$\mathcal{A}\left(v, m^d\right) := \underset{m}{\operatorname{argmin}} \frac{1}{2}\left(1 - m\right)^2 v + g\left(m, m^d\right) = \max\left(1 - \frac{1}{v}, m^d\right). \tag{75}$$

The following Proposition summarizes this.

Proposition 9.1 (Endogenizing attention) *The agent's attention m is given by:*

$$m = \mathcal{A}\left(\frac{\lambda\sigma_S^2}{\mathcal{K}}, m^d\right) = \max\left(1 - \frac{\mathcal{K}}{\lambda\sigma_S^2}, m^d\right),$$

where λ is in (74), $\mathcal{K}$ is the cost of cognition, and the attention function $\mathcal{A}$ is given by (75).

Proposition 9.1, illustrated by Figure 6, has a number of implications. First, *local rigidity from sparsity*: when σ_S^2 is low enough or $\mathcal{K}$ is large enough (so that $v = \frac{\lambda\sigma_S^2}{\mathcal{K}}$ is low and we are on the left, constant part of Figure 6), we have $m = m^d$, without any reaction to incentives. Locally, m is constant, at its default value m^d. Second, *flexible reaction of attention to strong incentives*: for large enough σ_S^2 (so that v is high and we are on the right, increasing part of Figure 6), then attention m increases in the variance, σ_S^2 and in the stakes λ. That is, the Lucas critique applies to big changes in parameters, but not to small ones.

This predicts for instance that when economic volatility (coming from TFP or monetary policy shocks) is increased, then: for a while, m does not move, and then, m increases as a function of volatility. For instance, people react more to interest rates in a highly volatile interest rate environment. This kind of comparative statics is sensible, and could be tested more systematically.

Concrete values for attention I now apply Proposition 9.1 to the consumer's attention, then to the firm's attention. I consider the case where all fluctuations are driven by productivity ζ_t, with the Taylor rule followed by monetary policy (I take $i_t = \phi_\pi \pi_t + \phi_x x_t + \bar{r}$, which is consistent with 0 inflation

41

on average, and this economy is worked out in Section 12.12). The proofs suggest how other sources of shocks could be handled.

Proposition 9.2 (Endogenizing the consumer's attention) *In the consumer problem, the attention is*

$$\bar{m} = \bar{m}^c = \mathcal{A}\left(\frac{\lambda^{\bar{m}}\sigma_\zeta^2}{\mathcal{K}^c}, \bar{m}^d\right), \qquad m_r = \mathcal{A}\left(\frac{\lambda^{m_r}\sigma_\zeta^2}{\mathcal{K}^c}, m_r^d\right), \qquad m_y = \mathcal{A}\left(\frac{\lambda^{m_y}\sigma_\zeta^2}{\mathcal{K}^c}, m_r^d\right), \tag{76}$$

with

$$(\lambda^{\bar{m}}, \lambda^{m_r}, \lambda^{m_y}) = \frac{\gamma(\gamma+\phi)}{\phi}\left(c_{\bar{m},\zeta}^2, c_{m_r,\zeta}^2, c_{m_y,\zeta}^2\right),$$

where $\mathcal{K}^c$ is the consumer's cost of cognition, the attention function $\mathcal{A}$ is in (75), and the coefficients $c_{\bar{m},\zeta}$, $c_{m_r,\zeta}$, $c_{m_y,\zeta}$ on the right-hand side are given in equations (224), (226) and (227) in the Online Appendix.

The intuition for those expressions is as follows. In this expression, $c_{m_r,\zeta}\zeta_t$ represents how much consumption changes (for a given ζ_t) if the consumer pays more attention to the interest rate. Hence, attention to the interest rate is higher if the interest rate matters more, i.e. $c_{m_r,\zeta}^2$ is high. So, when interest rates have moderate volatility, attention doesn't move (it stays at m_r^d), but when volatility increases much, then attention increases.

Proposition 9.3 (Endogenizing the firm's attention) *The firm's attention is:*

$$\bar{m} = \bar{m}^f = \mathcal{A}\left(\frac{\lambda^{\bar{m}^f}\sigma_\zeta^2}{\mathcal{K}^f}, \bar{m}^d\right), \qquad m_x^f = \mathcal{A}\left(\frac{\lambda^{m_x^f}\sigma_\zeta^2}{\mathcal{K}^f}, m_x^{f,d}\right), \qquad m_\pi^f = \mathcal{A}\left(\frac{\lambda^{m_\pi^f}\sigma_\zeta^2}{\mathcal{K}^f}, m_\pi^{f,d}\right), \tag{77}$$

with

$$\left(\lambda^f, \lambda^{m_x^f}, \lambda^{m_\pi^f}\right) = \frac{\varepsilon-1}{1-\beta\theta}\left(q_{\bar{m},\zeta}^2, q_{m_x^f,\zeta}^2, q_{m_\pi^f,\zeta}^2\right), \tag{78}$$

where $\mathcal{K}^f$ is the firms' cost of cognition, and the coefficients $q_{\bar{m},\zeta}$, $q_{m_x^f,\zeta}$, $q_{m_\pi^f,\zeta}$ on the right-hand size are given in equations (232), (235) and (236) in the Online Appendix.

Note that this might enrich the policy analysis – e.g. attention to inflation depends on the Fed's aggressiveness in controlling inflation and vice-versa. I do not pursue this here.

10 Appendix: Behavioral New Keynesian Macro in a Two-Period Economy

Here I present a two-period model that captures some of the basic features of the behavioral New Keynesian model. I recommend it for entrants to this literature, as everything is very clear with two periods.

It is similar to the model taught in undergraduate textbooks, but with rigorous microfoundations: it makes explicit the behavioral economics foundations of that undergraduate model. It highlights the complementarity between cognitive frictions and pricing frictions.

It is a useful model in its own right: to consider extensions and variants, I found it easiest to start with this two-period model.

Basic setup.   Utility is:

$$\sum_{t=0}^{1} \beta^t u\left(c_t, N_t\right) \text{ with } u\left(c, N\right) = \frac{c^{1-\gamma} - 1}{1 - \gamma} - \frac{N^{1+\phi}}{1 + \phi}.$$

As in Section 2.4, there is an economy consisting of a Dixit-Stiglitz continuum of firms with Calvo pricing frictions. Calling GDP Y_t, the aggregate production function is $Y_t = N_t$ and the aggregate resource constraint is:

$$\text{Resource constraint: } Y_t = c_t + G_t = N_t, \tag{79}$$

where G_t is real government consumption. The real wage is ω_t. Labor supply is frictionless, so the agent respects his first order condition: $\omega_t u_c + u_N = 0$, i.e.

$$\text{Labor supply: } N_t^{\phi} = \omega_t c_t^{-\gamma}. \tag{80}$$

The economy at time 1. Let us assume that the time-1 economy has flexible prices and no government consumption, but for simplicity labor supply is rigid at $N = 1$ (this is a technological constraint). The real wage must equal productivity, $\omega_t = 1$, and output is $y_1 = c_1 = 1$.

The economy at time 0. Now, consider the consumption demand at time 0, for the rational consumer. Taking for now personal income y_t as given, he solves $\max_{(c_t)_{t=0,1}} \sum_{t=0}^{1} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma}$ subject to $\sum_{t=0}^{1} \frac{c_t}{R_0^t} = y_0 + \frac{y_1}{R_0}$. That gives

$$c_0 = b\left(y_0 + \frac{y_1}{R_0}\right), \tag{81}$$

$$b := \frac{1}{1 + \beta},$$

with log utility.[103] Here b is the marginal propensity to consume (given the labor supply).[104]

Let us assume for now that the government does not issue any debt nor consumes. Then, aggregate income equals aggregate consumption: $y_t = c_t$. Hence,[105]

$$c_0 = b\left(c_0 + \frac{c_1}{R_0}\right), \tag{82}$$

which yields the Euler equation $\beta R_0 \frac{c_0}{c_1} = 1$. I use the consumption function formulation (82) rather than this Euler equation. Indeed, the consumption function is the formulation that generalizes well to behavioral agents.

Monetary policy is effective with sticky prices.   At time $t = 0$, a fraction θ of firms have sticky prices – their prices are pre-determined at a value we will call P_0^d (if prices are sticky, then $P_0^d = P_{-1}$, but we could have $P_0^d = P_{-1} e^{\pi_0^d}$, where π_0^d is an "automatic" price increase pre-programmed at time

---

[103]In the general case, $b := \frac{1}{1 + \beta^{\psi} R_0^{\psi - 1}}$, calling $\psi = \frac{1}{\gamma}$ the intertemporal elasticity of substitution (IES). In this section I just use $\psi = 1$.

[104]This is different from the more subtle MPC inclusive of labor supply movements, which is $\frac{\phi}{\gamma + \phi} \frac{1}{1+\beta}$ when evaluated at $c = N = 1$.

[105]The production subsidy by the government, designed to eliminate markup distortions, is paid for by lump-sum taxes. The consumer receives it in profits, then pays it in taxes, so that his total income is just labor income.

−1, not reactive to time-0 economic conditions, as in Mankiw and Reis (2002) or Section 6.3).[106] As in section 2.4, a corrective wage subsidy is assumed to be in place, so that there are no price distortions on average. Other firms freely optimize their price, and hence optimally choose a price

$$P_0^* = \omega_0 P_0,\tag{83}$$

where ω_0 is the real wage. Indeed, prices will be flexible at $t = 1$, so only current conditions matter for the optimal price. By (12), the aggregate price level is:

$$P_0 = \left(\theta\left(P_0^d\right)^{1-\varepsilon} + (1-\theta)\left(P_0^*\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}},\tag{84}$$

as a fraction θ of firms set the price P_0^d and a fraction $1 - \theta$ set the price P_0^*.

To solve the problem, there are 6 unknowns $(c_0, N_0, \omega_0, P_0, P_0^*, R_0)$ and 5 equations ((79)–(80) and (82)-(84)). What to do?

In the model with flexible prices ($\theta = 0$), this means that the price level P_0 is indeterminate (as in the basic Arrow-Debreu model). However, real variables are determinate: for instance, any solution yields $c_0 = N_0 = 1$.

In the model with sticky prices ($\theta > 0$), there is a one-dimensional continuum of real equilibria. It is the central bank that chooses the real equilibrium, by selecting the nominal interest rate, or equivalently here, by choosing the real interest rate R_0.[107] This is the great power of the central bank.

The behavioral consumer and fiscal policy. We can now consider the case where the consumer is behavioral. If his true income at time 1 is $y_1 = y_1^d + \widehat{y}_1$, he sees only $y_1^s = y_1^d + \bar{m}\widehat{y}_1$ for some $\bar{m} \in [0, 1]$, which is the attention to future income shocks ($\bar{m} = 1$ if the consumer is rational). Here the default is the frictionless case, $y_1^d = c_1 = Y_1 = 1$.

But now suppose that (81) becomes:

$$c_0 = b\left(y_0 + \frac{y_1^d + \bar{m}\hat{y}_1}{R_0}\right).\tag{85}$$

Suppose that the government consumes G_0 at 0, nothing at time 1, and makes a transfer $\mathcal{T}_t$ to the agents at times $t = 0, 1$. Call $d_0 = G_0 + \mathcal{T}_0$ the deficit at time 0. The government must pay its debt at the end of time 1, which yields the fiscal balance equation:

$$R_0 d_0 + \mathcal{T}_1 = 0.\tag{86}$$

The real income of a consumer at time 0 is

$$y_0 = c_0 + G_0 + \mathcal{T}_0 = c_0 + d_0.$$

Indeed, labor and profit income equal the sales of the firms, $c_0 + G_0$, plus the transfer from the government, $\mathcal{T}_0$. Income at time 1 is $y_1 = Y_1 + \mathcal{T}_1$: GDP, plus the transfer from the government.[108]

---

[106]This feature is not essential. The reader can imagine the case $\pi_0^d = 0$.

[107]The central bank chooses the nominal rate. Given equilibrium inflation, that allows it to choose the real rate (when there are pricing frictions).

[108]As we assumed that period 1 has frictionless pricing and no government consumption, we have $c_1 = Y_1 = 1$. If $d_0 > 0$, then the transfer $\mathcal{T}_1$ is negative. Agents use the proceeds of the time-0 government bonds to pay their taxes at time 1.

Hence, (85) gives:

$$c_0 = b \left(c_0 + d_0 + \frac{Y_1 + \bar{m}\mathcal{T}_1}{R_0} \right).$$

Using the fiscal balance equation (86) we have:

$$c_0 = b \left(c_0 + (1 - \bar{m}) d_0 + \frac{Y_1}{R_0} \right),$$

and solving for c_0:

$$c_0 = \frac{b}{1-b} \left((1 - \bar{m}) d_0 + \frac{Y_1}{R_0} \right). \tag{87}$$

We see how the "Keynesian multiplier" $\frac{b}{1-b}$ arises.

When consumers are fully attentive, $\bar{m} = 1$, and deficits do not matter in (87). However, take the case of behavioral consumers, $\bar{m} \in [0, 1)$. Consider a transfer by the government $\mathcal{T}_0$, with no government consumption, $G_0 = 0$. Equation (87) means that a positive transfer $d_0 = \mathcal{T}_0$ stimulates activity. If the government gives the agent $\mathcal{T}_0 > 0$ dollars at time 0, he does not fully see that they will be taken back (with interest) at time 1, so that this is awash. Hence, given $\frac{Y_1}{R_0}$, the consumer is tempted to consume more.

To see the full effect, when prices are not frictionless, we need to take a stance on monetary policy to determine R_0. Here, assume that the central bank does not change the interest rate R_0.[109] Then, (87) implies that GDP ($Y_0 = c_0 + G_0$) changes as:

$$\frac{dY_0}{d\mathcal{T}_0} = \frac{b}{1-b} (1 - \bar{m}). \tag{88}$$

With rational agents, $\bar{m} = 1$, and fiscal policy has no impact. With behavioral agents, $\bar{m} < 1$ and fiscal policy has an impact: the Keynesian multiplier $\frac{b}{1-b}$, times $(1 - \bar{m})$, a measure of deviation from full rationality. I record these results in the next proposition.

Proposition 10.1 *Suppose that we have (partially) sticky prices, and the central bank keeps the real interest rate constant. Then, a lump-sum transfer $\mathcal{T}_0$ from the government at time 0 creates an increase in GDP:*

$$\frac{dY_0}{d\mathcal{T}_0} = b_d \tag{89}$$

where

$$b_d := \frac{b}{1-b} (1 - \bar{m}), \tag{90}$$

where $b = \frac{1}{1+\beta}$ is the marginal propensity to consume, under log utility. Likewise, government spending G_0 has the multiplier:

$$\frac{dY_0}{dG_0} = 1 + b_d. \tag{91}$$

We see that $\frac{dY_0}{d\mathcal{T}_0} > 0$ and $\frac{dY_0}{dG_0} > 1$ if and only if consumers are non-Ricardian, $\bar{m} < 1$.

This proposition also announces a result on government spending, that I now derive. Consider an

---

[109]With flexible prices ($\theta = 0$), we still have $\omega_0 = 1$, hence we still have $c_0 = N_0 = 1$. Hence, the interest rate R_0 has to increase. Therefore, to obtain an effect of a government transfer, we need both monetary frictions (partially sticky prices) and cognitive frictions (partial failure of Ricardian equivalence).

increase in G_0, assuming a constant monetary policy (i.e., a constant real interest rate R_0 – alternatively, the central bank might choose to change rates).[110] Equation (87) gives $\frac{dc_0}{dG_0} = b_d$, so that GDP, $Y_0 = c_0 + G_0$, has a multiplier $\frac{dY_0}{dG_0} = 1 + b_d$.

When $\bar{m} = 1$ (Ricardian equivalence), a change in G_0 creates no change in c_0. Only labor demand N_0 increases, hence, via (80), the real wage increases, and inflation increases. GDP is $Y_0 = c_0 + G_0$, so that the multiplier $\frac{dY_0}{dG_0}$ is equal to 1.

However, when $\bar{m} < 1$ (so that Ricardian equivalence fails), the multiplier $\frac{dY_0}{dG_0}$ is greater than 1. This is due to the reason invoked in undergraduate textbooks: people feel richer, so they spend more, which creates more demand. Here, we can assert that with good conscience – provided we allow for behavioral consumers. In the fully non-Ricardian limit $\bar{m} = 0$, the fiscal multiplier is $\frac{dY_0}{dG_0} = 1 + b_d = 1 + \frac{b}{1-b} = \frac{1}{1-b}$, i.e.

$$\frac{dY_0}{dG_0} = \frac{1}{1-b} = \frac{1}{1 - MPC} \tag{92}$$

like in the undergraduate IS-LM model. Hence, we obtain a microfoundation for the undergraduate model, which is the limit of fully behavioral agents ($\bar{m} = 0$).

Without Ricardian equivalence, the government consumption multiplier is greater than 1.[111] Again, this relies on monetary policy being passive, in the sense of keeping a constant real rate R_0. If the real interest rate rises (as it would with frictionless pricing), then the multiplier would fall to a value less than 1.

Old vs. New Keynesian model: a mixture via bounded rationality. The above derivations show that the model is a mix of Old and New Keynesian models. Here, we do obtain a microfoundation for the Old Keynesian story (somewhat modified). We see what is needed: some form of non-Ricardian behavior (here via bounded rationality), and of sticky prices. This behavioral model allows for a simple (and I think realistic) mixture of the two ideas.

For completeness, I describe the behavior of realized inflation – the Phillips curve. I describe other features in Section 12.13.

The Phillips curve. Taking a log-linear approximation around $P_t = 1$, with $p_t = \ln P_t$, (84) becomes: $p_0 = \theta p_0^d + (1 - \theta) p_0^*$. Subtracting p_0 on both sides gives $0 = \theta \left(p_0^d - p_0 \right) + (1 - \theta) \left(p_0^* - p_0 \right)$, i.e.

$$p_0 - p_0^d = \frac{1 - \theta}{\theta} \left(p_0^* - p_0 \right).$$

Recall that $P_0^d = P_{-1} e^{\pi_0^d}$, so inflation is $\pi_0 = p_0 - p_{-1} = \left(p_0 - p_0^d \right) + \left(p_0^d - p_{-1} \right)$, i.e.

$$\pi_0 = \frac{1 - \theta}{\theta} \left(p_0^* - p_0 \right) + \pi_0^d. \tag{93}$$

Via (83),

$$p_0^* - p_0 = \hat{\omega}_0, \tag{94}$$

---

[110]See Woodford (2011) for an analysis with rational agents.

[111]This idea is known in the Old Keynesian literature. Mankiw and Weinzierl (2011) consider late in their paper non-Ricardian agents, and find indeed a multiplier greater than 1. But to do that they use two types of agents, which makes the analytics quite complicated when generalizing to a large number of periods. The methodology here generalizes well to static and dynamic contexts.

where $\hat{\omega}_0 = \frac{\omega_0 - \omega_0^*}{\omega_0^*}$ is the percentage deviation of the real wage from productivity, $\omega_0^* = 1$. Because $c_0 + G_0 = N_0$, the labor supply condition (80) implies:

$$\hat{\omega}_0 = \phi \hat{N}_0 + \gamma \hat{c}_0 = \phi \left(\hat{c}_0 + G_0 \right) + \gamma \hat{c}_0 = (\phi + \gamma) \left(\hat{c}_0 + b_g G_0 \right),$$

with $b_g := \frac{\phi}{\phi + \gamma}$.

Call "natural" an economy with frictionless pricing. In that economy, $\hat{\omega}_0^n = 0$, so that natural consumption $\hat{c}_0^n$ (measured as a deviation from baseline consumption $\bar{c} = 1$) satisfies $0 = (\phi + \gamma) \left(\hat{c}_0^n + b_g G_0 \right)$, i.e. $\hat{c}_0^n = -b_g G_0$. Therefore, defining the output gap x_0 as the difference between actual and natural output, i.e. between actual and natural consumption:

$$x_0 = \hat{c}_0 - \hat{c}_0^n = \hat{c}_0 + b_g G_0, \tag{95}$$

we have $\hat{\omega}_0 = (\phi + \gamma) x_0$. Hence (94) becomes $p_0^* - p_0 = (\phi + \gamma) x_0$, and (93) yields:

$$\text{Phillips curve: } \pi_0 = \kappa x_0 + \pi_0^d, \tag{96}$$

with $\kappa := \frac{1-\theta}{\theta} (\phi + \gamma)$. Hence, we obtain an elementary Phillips curve: increases in the output gap x_0 lead to inflation, above the automatic adjustment π_0^d. Finally, GDP's deviation from trend is:

$$\hat{Y}_0 = \hat{c}_0 + G_0 = x_0 + (1 - b_g) G_0.$$

To synthesize, we gather the results. Recall that $Y_0^d = 1$, while π_0 is the inflation between time -1 (the pre-time 0 price level) and time 0.[112] The deviations of (c_0, G_0) from trend are from the baseline of $(1, 0)$.

Proposition 10.2 (Two-period behavioral Keynesian model) *In this 2-period model, we have for time-0 output gap x_0 and inflation π_0:*

$$x_0 = b_g G_0 + b_d \hat{d}_0 - \sigma \left(i_0 - \mathbb{E}\pi_1 - \bar{r} \right) \text{ (IS curve)}, \tag{97}$$

$$\pi_0 = \kappa x_0 + \pi_0^d \text{ (Phillips curve)}, \tag{98}$$

while GDP satisfies:

$$\hat{Y}_0 = (1 - b_g) G_0 + x_0 = \hat{c}_0 + G_0, \tag{99}$$

where G_0 is government consumption as a share of baseline GDP, $\hat{d}_0$ the budget deficit, $b_g = \frac{\phi}{\phi + \gamma}$ is the output gap sensitivity to government spending, $b_d = \frac{b}{1-b} (1 - \bar{m})$ is the sensitivity to deficits, $b = \frac{1}{1+\beta}$ is the marginal propensity to consume (given labor income), $i_0 - \mathbb{E}\pi_1$ is the expected real interest rate between periods 0 and 1, $\sigma = \frac{1}{R} = \beta$ with log utility, and $\kappa = \frac{1-\theta}{\theta} (\phi + \gamma)$.

This completes the derivation of the 2-period Keynesian model. The Online Appendix (Section 12.13) contains complements, including a discounted Euler equation.

---

[112]If the agent perceived only part of the change in the real rate, replacing R_0 with $(1 - m_r) R_0^d + m_r R_0$ in (87), then the expression in (97) would be the same, replacing $\sigma = \frac{1}{R}$ with $\sigma = \frac{m_r}{R}$.

11 Appendix: Complements

11.1 Details of the Perception of Future Taxes

Here I flesh out the assumptions and results useful for the fiscal part of Section 5. First, we observe that iterating (44) gives $B_\tau = B_t + R\sum_{u=t}^{\tau-1} d_u$, so that the transfer at time τ, $\mathcal{T}_\tau = -\frac{r}{R}B_\tau + d_\tau$ is:

$$\mathcal{T}_\tau = -\frac{r}{R}B_t + \left(d_\tau - r\sum_{u=t}^{\tau-1} d_u\right). \tag{100}$$

Here I detail the formalism useful for the perception of future taxes. Call $\boldsymbol{Z}_\tau = (B_\tau, d_\tau, d_{\tau+1}, d_{\tau+2}, ...)'$ the state vector (more properly, the part of it that concerns deficits).

Under the rational model, $\boldsymbol{Z}_{\tau+1} = \boldsymbol{H}\boldsymbol{Z}_\tau$ for a matrix $\boldsymbol{H}$ characterized by: $(\boldsymbol{H}\boldsymbol{Z})(1) = \boldsymbol{Z}(1) + R\boldsymbol{Z}(2)$ and $(\boldsymbol{H}\boldsymbol{Z})(i) = \boldsymbol{Z}(i+1)$ for $i > 1$, where $\boldsymbol{Z}(i)$ is the i−th component of vector $\boldsymbol{Z}$. The true transfer at time τ is $\mathcal{T}_\tau = -\frac{r}{R}B_\tau + d_\tau = \mathcal{T}(\boldsymbol{Z}_\tau)$, where

$$\mathcal{T}(\boldsymbol{Z}) := \boldsymbol{e}^\mathcal{T}\boldsymbol{Z}, \qquad \boldsymbol{e}^\mathcal{T} := \left(-\frac{r}{R}, 1, 0, 0, ...\right)'. \tag{101}$$

We take a behavioral agent at time t. He forms a mental model of events and values at future dates $\tau \geq t$. Under his subjective model, the law of motion of vector $\boldsymbol{Z}_\tau$ is:

$$\boldsymbol{Z}_{\tau+1} - \boldsymbol{Z}_t^d = \bar{m}\boldsymbol{H}\left(\boldsymbol{Z}_\tau - \boldsymbol{Z}_t^d\right). \tag{102}$$

This is, the agent "anchors" future debt on the current debt captured by $\boldsymbol{Z}_t^d = (B_t, 0, 0, \ldots)$, he does only a partial adjustment for the future innovations – as captured by $\bar{m}$. This cognitive discounting implies, as in (10), that we have:

$$\mathbb{E}_t^{BR}\left[\mathcal{T}\left(\boldsymbol{Z}_\tau - \boldsymbol{Z}_t^d\right)\right] = \bar{m}^{\tau-t}\mathbb{E}_t\left[\mathcal{T}\left(\boldsymbol{Z}_\tau - \boldsymbol{Z}_t^d\right)\right] = \bar{m}^{\tau-t}\mathbb{E}_t\left[d_\tau - r\sum_{u=t}^{\tau-1} d_u\right]. \tag{103}$$

This gives the future taxes, as perceived by the agent at time t:

$$\mathbb{E}_t^{BR}\left[\mathcal{T}\left(\boldsymbol{Z}_\tau\right)\right] = -\frac{r}{R}B_t + \bar{m}^{\tau-t}\mathbb{E}_t\left[d_\tau - r\sum_{u=t}^{\tau-1} d_u\right]. \tag{104}$$

This reflects a partially rational consumer. Suppose that there are no future deficits. Given initial debt B_t, the consumer will see that it will have to be repaid: he accurately foresees the part $-\frac{r}{R}B_t$ in the perception of future deficits (104). However, he sees only dimly future deficits and their impact on future taxes. This is captured by the term $\bar{m}^{\tau-t}$.

The perceived law of motion for wealth (6) is extended to

$$k_{\tau+1} = \left(1 + \bar{r} + \hat{r}\left(\boldsymbol{X}_\tau^Z\right)\right)\left(k_\tau + \bar{y} + \hat{y}_\tau\left(N_\tau, \boldsymbol{X}_\tau^Z\right) + \mathcal{T}\left(\boldsymbol{X}_\tau^Z\right) - c_\tau\right), \tag{105}$$

where the state vector $\boldsymbol{X}_\tau^Z := (\boldsymbol{X}_\tau, \boldsymbol{Z}_\tau)$ is the extended state vector, comprising both the traditional $\boldsymbol{X}_\tau$ of the main model and the debt-related part $\boldsymbol{Z}_\tau$.

11.2 Additional Proofs

Proof of Lemma 2.4 *Notations.* The proof of this lemma and that of Proposition 2.5 follow the steps and notations of Galí (2015, Sections 3.2–3.3). I simplify matters by assuming constant returns to scale ($\alpha = 0$ in Galí's notations). So, the nominal marginal cost at $t + k$ is simply ψ_{t+k}, not $\psi_{t+k|t}$. When referring to, say, equation (11) of Chapter 3 in Galí (2015), I write "equation (G11)". I replace the coefficient of relative risk aversion (σ in his notations) by γ (as in $u'(c) = c^{-\gamma}$).

The high-level proof is as follows. If firms were rational, firms resetting their price would choose on average price of:

$$p_t^* = p_t + (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \, \mathbb{E}_t \left[\pi_{t+1} + \ldots + \pi_{t+k} - \mu_{t+k} \right]$$

as in Galí. Behavioral firms do this, but with cognitive discounting, so that

$$p_t^* = p_t + (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta\bar{m})^k \, \mathbb{E}_t \left[\pi_{t+1} + \ldots + \pi_{t+k} - \mu_{t+k} \right]$$

But this is a bit loose, so let us now turn to a real proof. If the firm were free to choose its real (log) price q_{it} freely at all dates t, it would choose price q_{it}^* maximizing (14), i.e. $q_i^*(\boldsymbol{X}_\tau) := \text{argmax}_{q_i} v(q_i, \boldsymbol{X}_\tau)$, i.e. $e^{q_{it}^*} = \frac{1 - \tau_f}{1 - \frac{1}{\varepsilon}} MC_t$. The subsidy $\tau_f = \frac{1}{\varepsilon}$ was chosen to eliminate the monopoly distortion on average.

As said before (15), we fix a date t and consider future dates $\tau \geq t$. For firms facing the Calvo pricing friction, we have, much as in the traditional model, that the price is the weighted average of future (perceived) optimal prices:[113]

$$q_{it} = (1 - \beta\theta) \sum_{\tau \geq t} (\beta\theta)^{\tau - t} \, \mathbb{E}_t^{BR} \left[q_i^*(\boldsymbol{X}_\tau) \right]$$

Given that Lemma 2.2 gives $\mathbb{E}_t^{BR}[q_i^*(\boldsymbol{X}_\tau)] = \bar{m}^{\tau - t} \mathbb{E}_t[q_i^*(\boldsymbol{X}_\tau)]$, we have:

$$q_{it} = (1 - \beta\theta) \sum_{\tau \geq t} (\beta\theta\bar{m})^{\tau - t} \, \mathbb{E}_t \left[q_i^*(\boldsymbol{X}_\tau) \right], \tag{106}$$

which is a behavioral counterpart to Galí's (G11).

Given (14), we have $v_q^0 (q_i^*(\boldsymbol{X}_\tau) - \Pi(\boldsymbol{X}_\tau), \mu(\boldsymbol{X}_\tau), c(\boldsymbol{X}_\tau)) = 0$, i.e., linearizing:

$$q_i^*(\boldsymbol{X}_\tau) = \Pi(\boldsymbol{X}_\tau) - \mu(\boldsymbol{X}_\tau) = \Pi_\tau - \mu_\tau = \pi_{t+1} + \cdots + \pi_\tau - \mu_\tau. \tag{107}$$

So, we have the following counterpart to the equation right before (G16):

$$q_{it} = (1 - \beta\theta) \sum_{\tau \geq t} (\beta\theta)^{\tau - t} \bar{m}^{\tau - t} \mathbb{E}_t \left[\pi_{t+1} + \cdots + \pi_\tau - \mu_\tau \right].$$

---

[113]The proof is as in the traditional model: the FOC of problem (16) is $\mathbb{E}_t^{BR} \sum_{\tau \geq t} (\beta\theta)^{\tau - t} v_{q_i}(q_{it}, \boldsymbol{X}_\tau) = 0$ and linearizing around $q_i^*(\boldsymbol{X}_\tau)$, the FOC is $\mathbb{E}_t^{BR} \sum_{\tau \geq t} (\beta\theta)^{\tau - t} v_{q_i q_i}(q_i^*(\boldsymbol{X}_\tau), \boldsymbol{X}_\tau) \cdot (q_{it} - q_i^*(\boldsymbol{X}_\tau)) = 0$. Taking the Taylor expansion around 0 disturbances so $q_i^*(\boldsymbol{X}_\tau)$ close to 0, the terms $v_{q_i q_i}(q_i^*(\boldsymbol{X}_\tau), \boldsymbol{X}_\tau)$ are approximately constant and equal to $v_{q_i q_i}(0, 0)$ up to first order terms, and the FOC is (up to second order terms) $\mathbb{E}_t^{BR} \sum_{\tau \geq t} (\beta\theta)^{\tau - t} \left(q_{it} - q_i^{*,BR}(\boldsymbol{X}_\tau) \right) = 0$, which gives $q_{it} = (1 - \beta\theta) \mathbb{E}_t \sum_{\tau \geq t} (\beta\theta)^{\tau - t} [q_i^*(\boldsymbol{X}_\tau)]$.

Proof of Proposition 2.5 Here I present a proof for both the basic case (Proposition 2.5) and the extended case (Proposition 6.4). In the proof I use the additional parameters m_x^f, m_π^f that are introduced in the extended model of Section 6.2 for fullness of exposition, although the reader who is referring to the baseline model of Section 2 can simply set $m_x^f = m_\pi^f = 1$ throughout.

Let us define

$$\rho := \beta\theta\bar{m}, \tag{108}$$

and calculate

$$H_t := \sum_{k\geq 1} \rho^k \left(\pi_{t+1} + ... + \pi_{t+k}\right) = \sum_{i\geq 1} \pi_{t+i} \sum_{k\geq i} \rho^k = \sum_{i\geq 1} \pi_{t+i} \frac{\rho^i}{1-\rho} = \frac{1}{1-\rho} \sum_{i\geq 0} \pi_{t+i}\rho^i 1_{i>0}.$$

Firms who can reset their price choose a price p_t^* given in (26) in the basic case, and (58) in the extended case:

$$
\begin{aligned}
p_t^* - p_t &= (1-\beta\theta) \sum_{k=0}^{\infty} \rho^k \mathbb{E}_t \left[m_\pi^f (\pi_{t+1} + ... + \pi_{t+k}) - m_x^f \mu_{t+k} \right] = (1-\beta\theta) \mathbb{E}_t \left[m_\pi^f H_t - \sum_{k=0}^{\infty} \rho^k m_x^f \mu_{t+k} \right] \\
&= (1-\beta\theta) \sum_{k\geq 0} \rho^k \mathbb{E}_t \left[\frac{m_\pi^f}{1-\rho} \pi_{t+k} 1_{k>0} - m_x^f \mu_{t+k} \right] \\
&= \sum_{k\geq 0} \rho^k \mathbb{E}_t \left[m'_{f\pi} \pi_{t+k} 1_{k>0} - \mu'_{t+k} \right], \tag{109}
\end{aligned}
$$

where $m'_{f\pi} := \frac{1-\beta\theta}{1-\rho} m_\pi^f$ and $\mu'_t := m_x^f (1-\beta\theta) \mu_t$.

Inflation. As a fraction $1-\theta$ of firms reset their price, starting from p_{t-1} on average:

$$
\begin{aligned}
\pi_t &= p_t - p_{t-1} = (1-\theta)(p_t^* - p_{t-1}) \\
&= (1-\theta)(p_t^* - p_t + p_t - p_{t-1}) = (1-\theta)(p_t^* - p_t + \pi_t). \\
\pi_t &= \frac{1-\theta}{\theta}(p_t^* - p_t). \tag{110}
\end{aligned}
$$

Plugging this in (109) gives:

$$\pi_t = \frac{1-\theta}{\theta} \sum_{k\geq 0} \rho^k \mathbb{E}_t \left[m'_{f\pi} \pi_{t+k} 1_{k>0} - \mu'_{t+k} \right]. \tag{111}$$

Next, I use the forward operator F ($Fy_t := y_{t+1}$), which allows me to evaluate infinite sums compactly, as in:

$$\sum_{k=0}^{\infty} \rho^k y_{t+k} = \sum_{k=0}^{\infty} \rho^k \left(F^k y_t \right) = \left(\sum_{k=0}^{\infty} \rho^k F^k \right) y_t = (1-\rho F)^{-1} y_t. \tag{112}$$

Rewriting (111) using F gives

$$\pi_t = \frac{1-\theta}{\theta} \mathbb{E}_t (1-\rho F)^{-1} \left(m'_{f\pi} \rho F \pi_t - \mu'_t \right).$$

Hence, multiplying by $1-\rho F$, we obtain the key equation (which is a behavioral version of (G17)):

$$\pi_t = \beta M^f \mathbb{E}_t [\pi_{t+1}] - \lambda\mu_t, \tag{113}$$

50

with

$$\beta^f := \beta M^f = \rho \left(1 + \frac{1-\theta}{\theta} m'_{f\pi} \right) = \beta \bar{m} \left(\theta + \frac{1-\beta\theta}{1-\beta\theta\bar{m}} m^f_\pi (1-\theta) \right), \tag{114}$$

$$M^f = \bar{m} \left(\theta + \frac{1-\beta\theta}{1-\beta\theta\bar{m}} m^f_\pi (1-\theta) \right), \tag{115}$$

$$\lambda = m^f_x \frac{1-\theta}{\theta} (1-\beta\theta). $$

The rest of the proof is as in Galí. The labor supply is still (19), $N^\phi_t = \omega_t c^{-\gamma}_t$, and as the resource constraint (up to second order terms coming from the price dispersion due to sticky prices) is $c_t = e^{\zeta_t} N_t$, $\omega_t = e^{-\phi\zeta_t} c^{\gamma+\phi}_t$, i.e. $\hat{\omega}_t = -\phi\zeta_t + (\gamma+\phi)\,\hat{c}_t$, so that with $\mu_t := \zeta_t - \hat{\omega}_t$,

$$\mu_t = (1+\phi)\,\zeta_t - (\gamma+\phi)\,\hat{c}_t.$$

Next, as in the "natural" economy without pricing frictions, $\mu^n_t = 0$:

$$0 = (1+\phi)\,\zeta_t - (\gamma+\phi)\,\hat{c}^n_t.$$

Hence, subtracting the two equations, and using $x_t = \hat{c}_t - \hat{c}^n_t$,

$$\mu_t = -\,(\gamma+\phi)\,x_t. \tag{116}$$

Plugging this into (113), we obtain the behavioral version of (G22):

$$\pi_t = \beta M^f \mathbb{E}_t\,[\pi_{t+1}] + \kappa x_t,$$

with $\kappa = \lambda\,(\gamma+\phi)$, i.e.

$$\kappa = \bar{\kappa} m^f_x, \tag{117}$$

$$\bar{\kappa} = \left(\frac{1}{\theta} - 1 \right)(1-\beta\theta)\,(\gamma+\phi). \tag{118}$$

Proof of Proposition 3.1 We will use the following well-known fact (Bullard and Mitra (2002)):

Proposition 11.1 (Roots in unit circle) *Consider the polynomial $p(x) = x^2 + ax + b$. Its two roots satisfy $|x| < 1$ if and only if: $|a| - 1 < b < 1$.*

We calculate $p(x) := \det(x\boldsymbol{I} - \boldsymbol{A}) = x^2 + ax + b$ with

$$a = -\frac{M + \beta^f + \kappa\sigma + \beta^f \sigma\phi_x}{D}, \quad b = \frac{M\beta^f}{D},$$

with $D = 1 + \sigma\,(\phi_x + \kappa\phi_\pi)$. Proposition 11.1 indicates that the equilibrium is determinate iff: $|a| - 1 < b < 1$. Given that we assume nonnegative coefficients ϕ, $b < 1$ and $a < 0$. Hence the criterion is: $1 + b + a > 0$, i.e. $p(1) > 0$. Calculations show that this is (35).

Proof of Proposition 3.2 Go back to (38), assuming the first best after the ZLB, so $z_T = 0$. Then,

$$z_0\,(T) = (A_{ZLB} - I)^{-1}(A_{ZLB}^T - I)\underline{b}.$$

When condition (36) fails, one of the eigenvalues of A_{ZLB} is greater than 1 in modulus. Then, $\lim_{T\to\infty}\|A_{ZLB}^T\underline{b}\| = \infty$ (it is easy to verify that $\underline{b}$ is not exactly the eigenvector corresponding to the root less than 1 in modulus). Hence, $\lim_{T\to\infty}\|z_0\,(T)\| = \infty$. Furthermore, this explosion is a recession: given that the entries of A_{ZLB} are positive, and those of $\underline{b}$ are negative, each of the terms in $\left(I + A_{ZLB} + ... + A_{ZLB}^{T-1}\right)\underline{b}$ is negative, hence $z_0\,(T)$ has unboundedly negative inflation and output gap.

When condition (36) holds, all roots of A_{ZLB} are less than 1 in modulus. Hence, $\lim_{T\to\infty}z_0\,(T) = -(A_{ZLB} - I)^{-1}\underline{b}$, a finite value.

Proof of Proposition 4.2 The Lagrangian is

$$L = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[-\frac{1}{2}\left(\pi_t^2 + \vartheta x_t^2\right) + \Xi_t\left(\beta M^f \pi_{t+1} + \kappa x_t - \pi_t + \nu_t\right)\right],$$

where Ξ_t are Lagrange multipliers. The first order conditions are: $L_{x_t} = 0$ and $L_{\pi_t} = 0$, which give respectively $-\vartheta x_t + \kappa \Xi_t = 0$ and $-\pi_t - \Xi_t + M^f \Xi_{t-1} = 0$, i.e. $\Xi_t = \frac{\vartheta}{\kappa}x_t$ and $\pi_t = \frac{-\vartheta}{\kappa}\left(x_t - M^f x_{t-1}\right)$.

Proof of Proposition 4.3 The central bank today takes its future actions as given, and chooses x_t, π_t, i_t to minimize today's loss $-\frac{1}{2}(\pi_t^2 + \vartheta x_t^2)$ subject to the behavioral IS equation and behavioral NK Phillips curve. This is equivalent to

$$\max_{\pi_t, x_t} -\frac{1}{2}\left(\pi_t^2 + \vartheta x_t^2\right) \qquad \text{subject to} \quad \pi_t = \beta M^f \mathbb{E}\pi_{t+1} + \kappa x_t + \nu_t,$$

and i_t can be read off the IS equation. Hence, the Lagrangian is simply:

$$L = -\frac{1}{2}\left(\pi_t^2 + \vartheta x_t^2\right) + \Xi\left(\beta M^f \mathbb{E}\pi_{t+1} + \kappa x_t + \nu_t - \pi_t\right).$$

The first order conditions are: $L_{x_t} = 0$ and $L_{\pi_t} = 0$, i.e. $-\vartheta x_t + \kappa\Xi = 0$ and $-\pi_t - \Xi = 0$, which together yields $\pi_t = -\frac{\vartheta}{\kappa}x_t$. The explicit value of i_t is in Section 13.

Proof of Proposition 6.2 First, we state a simple result.

Proposition 11.2 (Consumption given beliefs) *Consider an agent maximizing over (c_τ, N_τ) utility $U_t = \mathbb{E}_t^{BR}\sum_{\tau=t}^{\infty}\beta^{\tau-t}u\,(c_\tau, N_\tau)$ subject to the law of motion for wealth (50). Up to second order terms (and for small wealth k_0), consumption is:*

$$c_t = \frac{r}{R}k_t + \bar{y} + \mathbb{E}_t^{BR}\left[\sum_{\tau\geq t}\frac{1}{R^{\tau-t}}\left(b_r\hat{r}^{BR}\left(X_\tau\right) + b_y\hat{y}^{BR}\left(N_\tau, X_\tau\right)\right)\right], \qquad (119)$$

where expectations are taken under the agent's subjective model of the world, $b_r = \frac{-1}{\gamma R^2}, b_y = \frac{\bar{r}}{R}$, and N_τ labor supply at the perceived optimum. The chosen labor supply is given by $N_t^\phi = \hat{y}_N^{BR}\left(N_t, X_t\right)c_t^{-\gamma}$.

It is stated as a function of an endogenous labor supply, because this is the form that is most useful in some derivations (Section 12.2 develops the case that explicitly solves for labor supply). Versions of this proposition were proven a number of times with minor variants (e.g. Eusepi and Preston (2011); Woodford (2013); Gabaix (2016); Auclert (2018)), but for completeness let us derive it.[114]

Proof of Proposition 11.2. For simplicity, we take the deterministic case (as we consider first order Taylor expansions, the general case is the deterministic case where the path of variables are their expected values). The agent wants to maximize, over c_t and N_t:

$$\mathcal{L} = \sum_{t \geq 0} \beta^t u\left(c_t, N_t\right) + \lambda \left(k_0 + \sum_{t \geq 0} q_t \left(\bar{y} + \hat{y}^{BR}\left(N_t, \boldsymbol{X}_t\right) - c_t \right) \right), \tag{120}$$

where $q_t := 1/\prod_{\tau=0}^{t-1}\left(1 + \bar{r} + \hat{r}^{BR}\left(\boldsymbol{X}_\tau\right)\right)$. Here we consider the decision at time 0, which is just a normalization. Consider first the problem of optimizing $\mathcal{L}$ over c_t (taking the value of $y_t := \bar{y} + \hat{y}^{BR}\left(N_t, \boldsymbol{X}_t\right)$ as given), the FOC is $\beta^t c_t^{-\gamma} = \lambda q_t$, so $c_t = c_0 \left(\frac{\beta^t}{q_t}\right)^\psi$ (with $\psi := \frac{1}{\gamma}$), for some $c_0 = \lambda^{-\psi}$. With $\Omega := k_0 + \sum_{t \geq 0} q_t y_t$ the total perceived wealth, the perceived budget constraint is $\Omega = \sum_{t \geq 0} q_t c_t = c_0 \sum_{t \geq 0} \beta^{t\psi} q_t^{1-\psi}$. This gives consumption at decision time:

$$c_0 = \frac{k_0 + \sum_{t \geq 0} q_t y_t}{\sum_{t \geq 0} \beta^{t\psi} q_t^{1-\psi}}. \tag{121}$$

From there, let us see the impact of a small change in income. We have, at the default value of interest rates, $q_t = \beta^t$, so $\sum_{t \geq 0} \beta^{t\psi} q_t^{1-\psi} = \frac{1}{1-\beta}$, so $c_0 = (1-\beta)\left(k_0 + \sum_{t \geq 0} \beta^t y_t\right)$. This yields $b_y = 1 - \beta = \frac{r}{R}$ in (119). The impact of the interest rate on current consumptions is similar, though a little tedious (it is done in Section 13).

Next, for labor supply at time 0, the behavioral agent optimizes (120) given his perceived model of the world. This gives $\mathcal{L}_{N_0} = 0$, i.e. $-N_0^\phi + \lambda \hat{y}_N^{BR}\left(N_0, \boldsymbol{X}_0\right) = 0$. As we saw that $c_0^{-\gamma} = \lambda$, we obtain $N_0^\phi = c_0^{-\gamma} \hat{y}_N^{BR}\left(N_0, \boldsymbol{X}_0\right)$. $\square$

Application to this paper's behavioral agent *When $m_y = 1$, and no initial wealth.* This is the simplest case. The behavioral agent perceives his dynamic budget constraint is (50) and (51). Hence, we apply Proposition 11.2 to (51). At the optimum policy, we have $N_t = N\left(\boldsymbol{X}_t\right)$ under the perceived motion for $\boldsymbol{X}_t$, so the planned labor supply also verifies $N_\tau = N\left(\boldsymbol{X}_\tau\right)$, so $\hat{y}^{BR}\left(N_t, \boldsymbol{X}_t\right) = \hat{y}\left(\boldsymbol{X}_t\right)$. Now, using cognitive discounting (10), we have (52). This gives the consumption:

$$\hat{c}_t = \mathbb{E}_t^{BR}\left[\sum_{\tau \geq t} \frac{1}{R^{\tau-t}}\left(b_r \hat{r}^{BR}\left(\boldsymbol{X}_t\right) + \frac{r}{R} \hat{y}\left(\boldsymbol{X}_t\right)\right)\right] = \mathbb{E}_t\left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(b_r \hat{r}\left(\boldsymbol{X}_t\right) + \frac{r}{R} \hat{y}\left(\boldsymbol{X}_t\right)\right)\right].$$

With a general m_y or non-zero initial wealth. Here things are more complex, because the aggregate wealth is 0, and the agent plans to have non-zero wealth next period (as he misperceives income), so the agent doesn't plan that her future labor supply will be equal to the aggregate labor supply, $N_\tau = N\left(\boldsymbol{X}_\tau\right)$. Section 12.2 gives the proof.

Another method: Dynamic programming with Taylor expansion. The following method is a bit less intuitive, but may be handy to automatize when considering medium-scale extensions of this model.

---

[114]See also the proof of Lemma 4.2 in Gabaix (2016) for another style of derivation, using dynamic programming.

The subjective value function of the agent satisfies:

$$
\begin{aligned}
V(k, \boldsymbol{X}) \;=\; \max_{c,N}\{\, & u(c, N) \\
& + \beta \mathbb{E} V\left((1 + r + m_r \hat{r}(\boldsymbol{X}))(k + \bar{y} + m_y \hat{y}(\boldsymbol{X}) + w(\boldsymbol{X})(N - N(\boldsymbol{X})) - c), \bar{m}\boldsymbol{G^X}(\boldsymbol{X}, \boldsymbol{\epsilon})\right)\}
\end{aligned}
$$

and optimal consumption satisfies $u_c(c(k, \boldsymbol{X}), N) = V_k(k, \boldsymbol{X})$ (independently of N because utility is separable), so that $c_{\boldsymbol{X}} = \frac{V_{k\boldsymbol{X}}}{u_{cc}} = -\frac{V_{k\boldsymbol{X}}}{\gamma}$. In turn, $\hat{c}_t = c_{\boldsymbol{X}} \boldsymbol{X}_t$. Hence, to derive consumption, we simply need to calculate $V_{k\boldsymbol{X}}$. This is done in Section 12.15.

Proof of Proposition 6.3 Proposition 6.2 gives:

$$
\hat{c}_t = \mathbb{E}_t \left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}} \left(\frac{r}{R} m_Y \hat{y}_\tau + b_r m_r \hat{r}_\tau \right) \right]. \tag{122}
$$

Now, since there is no capital in the NK model, we have $\hat{y}_\tau = \hat{c}_\tau$: income is equal to aggregate demand. Hence, using $\tilde{b}_y := \frac{r}{R} m_Y$ and $\tilde{b}_r := b_r m_r = -\frac{m_r}{\gamma R^2}$, (122) becomes:

$$
\hat{c}_t = \mathbb{E}_t \left[\sum_{\tau \geq t} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}} \left(\tilde{b}_y \hat{c}_\tau + \tilde{b}_r \hat{r}_\tau \right) \right]. \tag{123}
$$

Taking out the first term yields:

$$
\hat{c}_t = \tilde{b}_y \hat{c}_t + \tilde{b}_r \hat{r}_t + \mathbb{E}_t \left[\sum_{\tau \geq t+1} \frac{\bar{m}^{\tau-t}}{R^{\tau-t}} \left(\tilde{b}_y \hat{c}_\tau + \tilde{b}_r \hat{r}_\tau \right) \right].
$$

Given that (123), applied to $t + 1$, yields $\hat{c}_{t+1} = \mathbb{E}_{t+1} \left[\sum_{\tau \geq t+1} \frac{\bar{m}^{\tau-t-1}}{R^{\tau-t-1}} \left(\tilde{b}_y \hat{c}_\tau + \tilde{b}_r \hat{r}_\tau \right) \right]$, we have:

$$
\hat{c}_t = \tilde{b}_y \hat{c}_t + \tilde{b}_r \hat{r}_t + \frac{\bar{m}}{R} \mathbb{E}_t [\hat{c}_{t+1}] = \frac{r}{R} m_Y \hat{c}_t + \tilde{b}_r \hat{r}_t + \frac{\bar{m}}{R} \mathbb{E}_t [\hat{c}_{t+1}].
$$

Multiplying by R and gathering the $\hat{c}_t$ terms, we have $\hat{c}_t = \frac{\bar{m}\mathbb{E}_t[\hat{c}_{t+1}] + R\tilde{b}_r \hat{r}_t}{R - rm_Y}$. This suggests defining $M := \frac{\bar{m}}{R - rm_Y}$ and $\sigma := \frac{-R\tilde{b}_r}{R - rm_Y} = \frac{m_r}{\gamma R(R - rm_Y)}$, and we get (18). This then translates into (24).[115]

Proof of Proposition 6.5 Call $q_{it} := p_{it} - p_t$ the real log price of firm i at a date t. Consider a firm that has not done a Calvo reset between t and $\tau > t$, and instead has simply passively indexed on default inflation. Then (using $\Delta z_\tau = z_\tau - z_{\tau-1}$), $\Delta p_{i,\tau} = \pi_\tau^d$, and $\Delta p_\tau = \pi_\tau = \pi_\tau^d + \hat{\pi}_\tau$. Hence, $\Delta q_{i,\tau} = -\hat{\pi}_\tau$, i.e. $q_{i\tau} = q_{it} - \hat{\Pi}_{i\tau}$ where $\hat{\Pi}_\tau := \hat{\pi}_{t+1} + \cdots + \hat{\pi}_\tau$ is the cumulative inflation between t and τ, but only in "hat space", i.e. considering the deviation of inflation from default inflation. Intuitively, the firm knows that it will indexed on default inflation, so it's important for it to forecast the deviation from default inflation, $\hat{\pi}_\tau$, not inflation itself.

---

[115]Here, bounded rationality lowers σ, the effective sensitivity to the interest rate, in addition to lowering M. With heterogeneous agents (along the lines of Auclert (2018)), one can imagine that bounded rationality might increase σ: some high-MPC (marginal propensity to consume) agents will have to pay adjustable-rate mortgages, which will increase the stimulative effects of a fall in the rate (increase σ).

Then, we are in a world isomorphic to that of Section 2.4, except that we put hats on π and Π – this is, replace π and Π by $\hat{\pi}$ and $\hat{\Pi}$. For instance, the state space is now $\boldsymbol{X}_\tau = \left(\boldsymbol{X}_\tau^{\mathcal{M}}, \hat{\Pi}_\tau\right)$, where $\boldsymbol{X}_\tau^{\mathcal{M}}$ is the basic macro state vector. The firm's profit is (14) with hats on Π, and so the natural generalization of (56) is that the behavioral firm perceives the flow profit

$$v^{BR}\left(q_{it}, \boldsymbol{X}_\tau\right) := v\left(q_{it} - m_\pi^f \hat{\Pi}\left(\boldsymbol{X}_\tau\right), m_x^f \mu\left(\boldsymbol{X}_\tau\right), c\left(\boldsymbol{X}_\tau\right)\right), \tag{124}$$

and its objective is still (16), with that notation. This leads to the economy in Proposition 6.5.[116]

References

Afrouzi, Hassan, "Strategic Inattention, Inflation Dynamics and the Non-Neutrality of Money," *Working Paper*, 2017.

Angeletos, George-Marios and Chen Lian, "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination," *Handbook of Macroeconomics*, 2016, *2*, 1065–1240.

_ **and** _ , "Dampening General Equilibrium: From Micro to Macro," *NBER Working Paper No. 22785*, 2017.

_ **and** _ , "Forward guidance without common knowledge," *American Economic Review*, 2018, *108* (9), 2477–2512.

_ **and Jennifer La'O**, "Dispersed information over the business cycle: optimal fiscal and monetary policy," *Working Paper*, 2008.

_ **and** _ , "Incomplete Information, Higher-Order Beliefs and Price Inertia," *Journal of Monetary Economics*, 2009, *56*, S19–S37.

_ **and** _ , "Noisy Business Cycles," *NBER Macroeconomics Annual*, 2010, *24* (1), 319–378.

_ **and** _ , "Optimal monetary policy with informational frictions," *Forthcoming at the Journal of Political Economy*, 2018.

_ **and Zhen Huo**, "Myopia and Anchoring," *Working Paper*, 2019.

_ **, Fabrice Collard, and Harris Dellas**, "Quantifying confidence," *Econometrica*, 2018, *86* (5), 1689–1726.

Ascari, Guido and Tiziano Ropele, "Sacrifice Ratio in a Medium-Scale New Keynesian Model," *Journal of Money, Credit and Banking*, 2012, *44* (2-3), 457–467.

Auclert, Adrien, "Monetary Policy and the Redistribution Channel," *Forthcoming at the American Economic Review*, 2018.

Barro, Robert J, "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 1974, *82* (6), 1095–1117.

Benchimol, Jonathan and Lachen Bounader, "Optimal Monetary Policy Under Bounded Rationality," *Working Paper*, 2018.

Bilbiie, Florin O, "Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic," *Journal of Economic Theory*, 2008, *140* (1), 162–196.

Bilbiie, Florin Ovidiu, "Monetary Policy and Heterogeneity: An Analytical Framework," *Working Paper*, 2018.

---

[116]The derivation is very simple. Lemma 2.4 holds, putting hats on π, and as a result, the Phillips curve holds (28) also putting hats on π, i.e. (61) holds.

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

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," *Journal of Political Economy*, 2013, *121* (5), 803–843.

_ , _ , **Yueran Ma, and Andrei Shleifer**, "Over-reaction in macroeconomic expectations," *NBER Working Paper No. 24932*, 2018.

Bounader, Lahcen, "Optimal Monetary Policy in Behavioral New Keynesian Model," *Working Paper*, 2016.

Brown, Jennifer, Tanjim Hossain, and John Morgan, "Shrouded Attributes and Information Suppression: Evidence from the Field," *Quarterly Journal of Economics*, 2010, *125* (2), 859–876.

Bullard, James and Kaushik Mitra, "Learning about monetary policy rules," *Journal of Monetary Economics*, 2002, *49* (6), 1105–1129.

Caballero, Ricardo J and Emmanuel Farhi, "The Safety Trap," *Review of Economic Studies*, 2017, *85* (1), 223–274.

Campbell, Jeffrey R, Jonas DM Fisher, Alejandro Justiniano, and Leonardo Melosi, "Forward Guidance and Macroeconomic Outcomes since the Financial Crisis," *NBER Macroeconomics Annual*, 2017, *31* (1), 283–357.

Campbell, John Y. and N. Gregory Mankiw, "Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence," *NBER Macroeconomics Annual*, 1989, *4*, 185–216.

Caplin, Andrew, Mark Dean, and Daniel Martin, "Search and Satisficing," *The American Economic Review*, 2011, *101* (7), 2899–2922.

_ , _ , **and John Leahy**, "Rationally inattentive behavior: Characterizing and generalizing Shannon entropy," *NBER Working Paper No. 23652*, 2017.

Carvalho, Carlos and Fernanda Nechio, "Do People Understand Monetary Policy?," *Journal of Monetary Economics*, 2014, *66*, 108–123.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 1–35.

Châu, Vũ T, "Behavioral-attention Phillips curve: Theory and evidence from inflation surveys," *Working Paper*, 2019.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and Taxation: Theory and Evidence," *The American Economic Review*, 2009, *99* (4), 1145–1177.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Nominal rigidities and the dynamic effects of a shock to monetary policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

Clarida, Richard, Jordi Gali, and Mark Gertler, "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 1999, *37* (4), 1661–1707.

_ , _ , **and** _ , "Monetary policy rules and macroeconomic stability: evidence and some theory," *The Quarterly Journal of Economics*, 2000, *115* (1), 147–180.

Cochrane, John H, "Determinacy and Identification with Taylor Rules," *Journal of Political Economy*, 2011, *119* (3), 565–615.

_ , "Michelson-Morley, Occam and Fisher: The Radical Implications of Stable Inflation at Near-Zero Interest Rates," *NBER Macroeconomics Annual*, 2017, *32* (1), 113–126.

Coibion, Olivier and Yuriy Gorodnichenko, "What can survey forecasts tell us about information rigidities?," *Journal of Political Economy*, 2012, *120* (1), 116–159.

_ **and** _ , "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *The American Economic Review*, 2015, *105* (8), 2644–2678.

Correia, Isabel, Juan Pablo Nicolini, and Pedro Teles, "Optimal fiscal and monetary policy: Equivalence results," *Journal of Political Economy*, 2008, *116* (1), 141–170.

Eggertsson, Gauti B, "What fiscal policy is effective at zero interest rates?," *NBER Macroeconomics Annual*, 2011, *25* (1), 59–112.

Eggertsson, Gauti B. and Michael Woodford, "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 2003, *34* (1), 139–235.

Eggertsson, Gauti B and Paul Krugman, "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach," *Quarterly Journal of Economics*, 2012, *127* (3), 1469–1513.

_ **, Neil R Mehrotra, and Jacob A Robbins**, "A model of secular stagnation: Theory and quantitative evaluation," *American Economic Journal: Macroeconomics*, 2019, *11* (1), 1–48.

Erceg, Christopher J and Andrew T Levin, "Imperfect credibility and inflation persistence," *Journal of Monetary Economics*, 2003, *50* (4), 915–944.

Eusepi, Stefano and Bruce Preston, "Expectations, learning, and business cycle fluctuations," *The American Economic Review*, 2011, *101* (6), 2844–2872.

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

Evans, George W and Seppo Honkapohja, *Learning and Expectations in Macroeconomics*, Princeton University Press, 2001.

Eyster, Erik, Kristof Madarasz, and Pascal Michaillat, "Pricing when Customers Care about Fairness but Misinfer Markups," *NBER Working Paper No. 23778*, 2017.

Farhi, Emmanuel and Iván Werning, "Monetary Policy, Bounded Rationality, and Incomplete Markets," *NBER Working Paper No. 23281*, 2017.

_ **and Xavier Gabaix**, "Optimal Taxation with Behavioral Agents," *NBER Working Paper No. 21524*, 2017.

Fisher, Jonas DM, "On the Structural Interpretation of the Smets-Wouters 'Risk Premium' Shock," *Journal of Money, Credit and Banking*, 2015, *47* (2-3), 511–516.

Fuhrer, Jeff, "Expectations as a source of macroeconomic persistence: Evidence from survey expectations in a dynamic macro model," *Journal of Monetary Economics*, 2017, *86*, 22–35.

Fuhrer, Jeffrey C and Glenn D Rudebusch, "Estimating the Euler equation for output," *Journal of Monetary Economics*, 2004, *51* (6), 1133–1153.

Fuster, Andreas, Benjamin Hebert, and David Laibson, "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing," *NBER Macroeconomics Annual*, 2012, *26* (1), 1–48.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 2014, *129* (4), 1661–1710.

_ , "Behavioral Macroeconomics via Sparse Dynamic Programming," *NBER Working Paper No. 21848*, 2016.

_ , "Behavioral Inattention," *Handbook of Behavioral Economics*, 2019, pp. 261–344.

_ **and David Laibson**, "The 6D Bias and the Equity-Premium Puzzle," *NBER Macroeconomics Annual*, 2002, *16*, 257–312.

_ **and** _ , "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

_ **and** _ , "Myopia and Discounting," *NBER Working Paper No. 23254*, 2017.

Galí, Jordi, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and its Applications*, Princeton University Press, 2015.

_ , "Monetary Policy and Bubbles in a New Keynesian Model with Overlapping Generations," *Working Paper*, Mar 2017.

_ **and Mark Gertler**, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 1999, *44* (2), 195–222.

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

Ganong, Peter and Pascal J Noel, "Consumer spending during unemployment: Positive and normative implications," *Forthcoming at the American Economic Review*, 2019.

García-Schmidt, Mariana and Michael Woodford, "Are low interest rates deflationary? A paradox of perfect-foresight analysis," *American Economic Review*, 2019, *109* (1), 86–120.

Greenwood, Robin and Andrei Shleifer, "Expectations of returns and expected returns," *The Review of Financial Studies*, 2014, *27* (3), 714–746.

_ **and Samuel G. Hanson**, "Waves in ship prices and investment," *Quarterly Journal of Economics*, 2015, *130* (1), 55–109.

Hébert, Benjamin and Michael Woodford, "Information costs and sequential information sampling," *NBER Working Paper No. 25316*, 2018.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *The American Economic Review*, December 2006, *96* (5), 1589–1610.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante, "Monetary policy according to HANK," *American Economic Review*, 2018, *108* (3), 697–743.

Kiley, Michael T, "Endogenous price stickiness and business cycle persistence," *Journal of Money, Credit and Banking*, 2000, pp. 28–53.

_ , "Policy Paradoxes in the New Keynesian Model," *Review of Economic Dynamics*, 2016, *21*, 1–15.

Kocherlakota, Narayana, "Fragility of Purely Real Macroeconomic Models," *NBER Working Paper No. 21866*, 2016.

Lindé, Jesper, "Estimating New-Keynesian Phillips curves: A full information maximum likelihood approach," *Journal of Monetary Economics*, 2005, *52* (6), 1135–1149.

Maćkowiak, Bartosz and Mirko Wiederholt, "Business Cycle Dynamics Under Rational Inattention," *Review of Economic Studies*, 2015, *82* (4), 1502–1532.

Malmendier, Ulrike and Stefan Nagel, "Learning from inflation experiences," *The Quarterly Journal of Economics*, 2015, *131* (1), 53–87.

Mankiw, N. Gregory, "The Savers-Spenders Theory of Fiscal Policy," *The American Economic Review Papers and Proceedings*, 2000, *90* (2), 120–125.

_ **and Matthew Weinzierl**, "An Exploration of Optimal Stabilization Policy," *Brookings Papers on Economic Activity*, 2011, *42* (1 (Spring)), 209–272.

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

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H Stock, "Empirical evidence on inflation expectations in the New Keynesian Phillips Curve," *Journal of Economic Literature*, 2014, *52* (1), 124–188.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson, "The Power of Forward Guidance Revis-

ited," *The American Economic Review*, 2016, *106* (10), 3133–3158.

_ , _ , **and** _ , "The discounted Euler equation: A note," *Economica*, 2017, *84* (336), 820–831.

Morris, Stephen and Hyun Song Shin, "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *The American Economic Review*, 1998, pp. 587–597.

_ , **Franklin Allen, and Hyun Song Shin**, "Beauty Contests, Bubbles and Iterated Expectations in Asset Markets," *The Review of Financial Studies*, 2006, *19* (4), 719–752.

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

Nakata, Taisuke, Ryota Ogaki, Sebastian Schmidt, and Paul Yoo, "Attenuating the forward guidance puzzle: implications for optimal monetary policy," *Forthcoming at the Journal of Economic Dynamics and Control*, 2019.

Negro, Marco Del, Marc P Giannoni, and Christina Patterson, "The Forward Guidance Puzzle," *Federal Reserve Bank of New York Staff Reports 574*, 2015.

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

Paciello, Luigi and Mirko Wiederholt, "Exogenous information, endogenous information, and optimal monetary policy," *Review of Economic Studies*, 2013, *81* (1), 356–388.

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

Smets, Frank and Rafael Wouters, "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *The American Economic Review*, 2007, *97* (3), 586–606.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention variation and welfare: theory and evidence from a tax salience experiment," *The Review of Economic Studies*, 2017, *85* (4), 2462–2496.

Taylor, John B, "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank," *Journal of Monetary Economics*, 1999, *43* (3), 655–679.

Werning, Iván, "Managing a Liquidity Trap: Monetary and Fiscal Policy," *NBER Working Paper 17344*, 2012.

_ , "Incomplete Markets and Aggregate Demand," *NBER Working Paper No. 21448*, August 2015.

Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," in Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford, eds., *Knowledge, information, and expectations in modern macroeconomics: In honor of Edmund S. Phelps*, Princeton University Press, 2003.

_ , *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2003.

_ , "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 2011, *3* (1), 1–35.

_ , "Macroeconomic Analysis without the Rational Expectations Hypothesis," *Annual Reviews of Economics*, 2013, *5*, 303–346.

_ , "Monetary policy analysis when planning horizons are finite," *NBER Macroeconomics Annual*, 2018.