NBER WORKING PAPER SERIES

A BEHAVIORAL NEW KEYNESIAN MODEL

Xavier Gabaix

Working Paper 22954
http://www.nber.org/papers/w22954

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016

I thank Igor Cesarec, Vu Chau, Wu Di and James Graham for excellent research assistance. For useful comments I thank Marios Angeletos, Adrien Auclert, Jeff Campbell, Larry Christiano, John Cochrane, Gauti Eggertsson, Emmanuel Farhi, Jordi Galí, Mark Gertler, Narayana Kocherlakota, Ricardo Reis, Michael Woodford, and participants at various seminars and conferences. I am grateful to the CGEB, the Institute for New Economic Thinking, the NSF (SES-1325181) and the Sloan Foundation for financial support. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

A Behavioral New Keynesian Model
Xavier Gabaix
NBER Working Paper No. 22954
December 2016
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 parameter, which quantifies how poorly agents understand future policy and its impact. 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 reduce 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 helps solve 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 forms. (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 the empirical evidence.

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

A online appendix is available at http://www.nber.org/data-appendix/w22954

1 Introduction

This paper proposes a way to analyze what happens to 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 in 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 agents demands "nominal GDP 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 the 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.

[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.

Let me expand on the above points.

Fiscal policy and helicopter drops of money. In the traditional NK model, agents are fully rational. So Ricardian equivalence holds, and fiscal policy (i.e. 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, 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 a rational model, probably counterfactually so (e.g. 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.

The Taylor principle reconsidered and equilibrium determinacy. When monetary policy is passive (e.g. constant interest rate, 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.[2] In contrast, in this behavioral model there is just one (bounded) equilibrium: things are clean and definite theoretically.

Economic stability. Determinacy is not just a purely theoretical question. In the rational model, if the economy is stuck at the ZLB forever (or, I will argue, for a long period of time), the Taylor principle is violated (as the 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. However, we do not see that in Japan since the late 1980s or in the Western economies in the aftermath of the 2008 crisis (Cochrane 2015). 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, Giannoni and Patterson (2015) and McKay, Nakamura and Steinsson (2016). The reason is again that the traditional

[2]However, I do need to rule out explosive equilibria.

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 I put forth, this impact is muted by the agent's myopia, which makes forward guidance less powerful. The model, in reduced form, takes the form of a "discounted Euler equation", where the agent reacts in a discounted manner to future consumption growth.

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.[3] 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 great deflation and come back to the initial price level. 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 (2015), highlight that in the strict (rational) NK model, a rise in interest rates (even temporary) leads to a rise in inflation (though this depends on which equilibrium is selected, leading to some cacophony in the dialogue), which is paradoxical. This is called the "neo-Fisherian" property. In the present behavioral model, the property holds in the long run: the long-run real rate is independent of monetary policy (Fisher neutrality holds in that sense). However, in the short run, raising rates does lower inflation and output, as in the Keynesian model.

Literature review. I build on the large New Keynesian literature, as distilled in Woodford (2003) and Galí (2015). I am also indebted to the number of authors who identified paradoxes in the New Keynesian model, e.g. Cochrane (2015), Del Negro, Giannoni and Patterson (2015), McKay, Nakamura and Steinsson (2016).

For the behavioral model, I rely on the general dynamic setup derived in Gabaix (2016), itself building on a general static "sparsity approach" to behavioral economics laid out in Gabaix (2014). There are other ways to model bounded rationality, including related sorts of differential salience (Bordalo, Gennaioli and Shleifer 2016), rules of thumb (Campbell and Mankiw 1989), limited information updating (Caballero 1995, Gabaix and Laibson 2002, Mankiw and Reis 2002, Reis 2006),

[3] See Clarida, Galí and Gertler (1999).

noisy signals (Sims 2003, Maćkowiak and Wiederholt 2015, Woodford 2012).[4] The sparsity approach aims at being tractable and fairly unified, as it applies to both microeconomic problems like basic consumer theory and Arrow-Debreu-style general equilibrium (Gabaix 2014), dynamic macroeconomics (Gabaix 2016) and public economics (Farhi and Gabaix 2015). Its tractability comes in part from the fact that it does not rely on the extraction of noisy signals, which complicates other models.

At any rate, this is the first paper to study how behavioral considerations affect forward guidance in the New Keynesian model, alongside Garcia-Schmidt and Woodford (2015).[5] That paper was circulated simultaneously and offers very different modelling. Woodford (2013) explores non-rational expectations in the NK model, particularly of the learning type. However, he does not distill his rich analysis into something compact like the 2-equation NK model of Proposition 2.5 in this paper.

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 and fiscal policy with behavioral agents. Section 5 considers a slightly more complex version that allows the model to handle long-run changes. Section 6 concludes. Section 7 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 precisions.

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 (2003). 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 subtitution ($\sigma = 1/\gamma$ in the traditional model).

I will call the economy "determinate" (in the sense of Blanchard and Kahn (1980)) if there is only one non-explosive equilibrium path.

[4] My notion of "behavioral" here is bounded rationality or cognitive myopia. I abstract from other interesting forces, like fairness (Eyster, Madarasz and Michaillat 2016) – they create an additional source of price stickiness.

[5] The core model in this paper was circulated in the summer of 2015, as a part of what became Gabaix (2016). Works circulated a year later pursue related themes. Farhi and Werning (2016) explore the interaction between bounded rationality and heterogeneous agents (rather than a representative agent as in this paper). Angeletos and Lian (2016a,b) explore the difference between direct and general equilibrium contexts, with an application to forward guidance. Their notion of "incomplete information" is a close cousin of the limited attention I study here, with a different but related microfoundation.

2 A Behavioral Model

As in the traditional NK model, there is no capital or government spending, so output equals consumption, c_t. I call r_t^n (respectively, c_t^n) the "natural real interest rate" (respectively, natural consumption), which is the interest rate that would prevail if all pricing frictions were removed (but cognitive frictions are not eliminated).[6] I call $x_t = (c_t - c_t^n)/c_t^n$ the output gap. Hence, positive x_t corresponds to a boom, negative x_t to a recession. With rational agents, the traditional NK model gives microfoundations that lead to:

$$x_t = \mathbb{E}_t [x_{t+1}] - \frac{1}{\gamma} (i_t - \mathbb{E}_t \pi_{t+1} - r_t^n) \quad (1)$$

$$\pi_t = \beta \mathbb{E}_t [\pi_{t+1}] + \kappa x_t. \quad (2)$$

I now present other foundations leading to a behavioral model that has the traditional rational model as a particular case.

2.1 Behavioral Agent: Microeconomic Behavior

In this section I present a way of modeling boundedly rational agents, drawing results from Gabaix (2016), which discusses more general sparse behavioral dynamic programming.

Setup: Objective reality. I consider an agent with standard utility

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

where c_t is consumption, and N_t is labor supply (as in Number of hours supplied). The real wage is ω_t. The real interest rate is r_t and agent's real income is $y_t = \omega_t N_t + \Pi_t + \mathcal{T}_t$: the sum of labor income $\omega_t N_t$, profit income Π_t, and potential government transfers $\mathcal{T}_t$ (normally set to 0). His real financial wealth k_t evolves as:[7]

$$k_{t+1} = (1 + r_t)(k_t - c_t + y_t) \quad (3)$$

[6] More precisely, this is the interest rate in a surrogate economy with no pricing frictions, the first best level of consumption and labor supply, but with the same cognitive frictions as in the original economy, and before government transfers to the agents. Section 9.1 details this.

[7] I change a bit the timing convention compared to Gabaix (2016): income innovation $\hat{y}_t$ is received at t, not $t+1$. The wealth k_t is the wealth at the beginning of the period, before receiving any income.

The agent's problem is $\max_{(c_t,N_t)_{t\geq 0}} U$ s.t. (3), and the usual transversality condition.

The aggregate production of the economy is $c_t = e^{\zeta_t} N_t$, where ζ_t follows an AR(1). There is no capital, as in the baseline New Keynesian model. Hence, in most expressions $k_t = 0$ in equilibrium, but we need to consider potential deviations from $k_t = 0$ when studying the agent's consumption problem.

Consider first the case where the economy is deterministic at the steady state ($\zeta_t \equiv 0$), so that the interest rate, income, and real wage are at $\bar{r}$ (which I will soon call r), $\bar{y}$, and $\overline{\omega}$. We have a simple deterministic problem.[8] Defining $R := 1 + \bar{r}$, we have $\beta R = 1$. If there are no taxes and profits, as in the baseline model, $\bar{y} = \bar{\omega}\bar{N}$.

In general, there will be deviations from the steady state. I decompose the values as deviations from the above steady state:

$$r_t = \bar{r} + \hat{r}_t,\ y_t = \bar{y} + \hat{y}_t,\ \omega_t = \bar{\omega} + \hat{\omega}_t,\ N_t = \bar{N} + \hat{N}_t.$$

Also, there is a state vector $\boldsymbol{X}_t$ (comprising TFP ζ_t, announced actions in monetary and fiscal policy), that will evolve in equilibrium as:

$$\boldsymbol{X}_{t+1} = \boldsymbol{G}\left(\boldsymbol{X}_t, \boldsymbol{\epsilon}_{t+1}\right) \tag{4}$$

for some equilibrium transition function G and mean-0 innovations $\boldsymbol{\epsilon}_{t+1}$. I assume that $\boldsymbol{X}_t$ has mean 0, i.e. has been de-meaned. Linearizing, the law of motion becomes:

$$\boldsymbol{X}_{t+1} = \boldsymbol{\Gamma}\boldsymbol{X}_t + \boldsymbol{\epsilon}_{t+1} \tag{5}$$

for some matrix $\boldsymbol{\Gamma}$, after perhaps a renormalization of $\boldsymbol{\epsilon}_{t+1}$. In equilibrium, the deviation of the interest rate and income will depend on $\boldsymbol{X}_t$: linearizing, as $\hat{y}_t = \boldsymbol{b}^y_{\boldsymbol{X}}\boldsymbol{X}_t$, for some factor $\boldsymbol{b}^y_{\boldsymbol{X}}$.

[8] The solution is: $\bar{c} = \bar{N} = \bar{\omega} = 1$. 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: $\bar{c} = \bar{N}$.

Setup: Reality perceived by the behavioral agent. I can now describe the behavioral agent. I posit that he perceives that his wealth and the state vector evolve as:[9]

$$k_{t+1} = (1+\bar{r}+m_r\widehat{r}_t)(k_t+\bar{y}+m_y\hat{y}_t-c_t) \tag{6}$$

$$\boldsymbol{X}_{t+1} = \bar{m}\boldsymbol{G}(\boldsymbol{X}_t,\boldsymbol{\epsilon}_{t+1}) \tag{7}$$

where $m_r, m_y, \bar{m}$ are attention parameters in $[0,1]$. When they are all equal to 1, the agent is the traditional, rational agent. Here m_r and m_y capture the 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. To better interpret $\bar{m}$, we linearize (7):

$$\boldsymbol{X}_{t+1} = \bar{m}(\boldsymbol{\Gamma}\boldsymbol{X}_t + \boldsymbol{\epsilon}_{t+1}) \tag{8}$$

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

$$\mathbb{E}_t^{BR}[\boldsymbol{X}_{t+k}] = \bar{m}^k\mathbb{E}_t[\boldsymbol{X}_{t+k}] \tag{9}$$

where the $\mathbb{E}_t^{BR}[\boldsymbol{X}_{t+k}]$ is the subjective expectation by the behavioral agent, and $\mathbb{E}_t[\boldsymbol{X}_{t+k}]$ is the rational expectations. 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 $\bar{m} \in [0,1]$). Parameter $\bar{m}$ is a form of "global cognitive discounting" – discounting future disturbances more as they are more distant in the future.[10]

Gabaix (2014) discusses the microfoundation of the inattention parameters, and proposes an endogenization for them.[11] To summarize, they can come from costs to attention, so that an agent will choose limited attention. Or, each agent could get noisy signals, and then as in Bayesian updating his average perception will be a dampened value of the true parameter. This "sparsity"

[9]This is in the spirit of Gabaix (2016). In that paper, I didn't explore the "cognitive discounting" $\bar{m}$, but it is admissible by the general formalism.

[10]See Gabaix (2016), Section 12.10 in the online appendix.

[11]In this paper, though, I will keep the parameters invariant to policy. A deeper endogenization would make them depend on policy, which would be an interesting extension.

framework formalizes those things, while keeping a tractable model that applies beyond the usual Gaussian-linear-quadratic context. This said, in this paper, I insist on features that do not depend on the fine details of how attention is modeled.

Term structure of attention to income and interest rate. Recall that income is perceived with an attention factor m_y (see (6)) and that $\hat{y}_t = \boldsymbol{b}_y^X \boldsymbol{X}_t$, for some factor $\boldsymbol{b}_y^X$. This implies:

$$\mathbb{E}_t^{BR}[\hat{y}_{t+k}] = m_y \bar{m}^k \mathbb{E}_t[\hat{y}_{t+k}], \tag{10}$$

$$\mathbb{E}_t^{BR}[\hat{r}_{t+k}] = m_r \bar{m}^k \mathbb{E}_t[\hat{r}_{t+k}]. \tag{11}$$

Hence, we obtain a "term structure of attention". Factor m_y is the "level" or "intercept" of attention, while factor $\bar{m}$ is the "slope" of attention as a function of the horizon.

There is mounting microeconomic evidence for the existence of inattention to macroeconomic variables (Coibon and Gorodnichenko 2015), taxes (Chetty, Looney and Kroft 2009, Taubinsky and Rees-Jones 2015), and, more generally, small dimensions of reality (Brown, Hossain and Morgan 2010, Caplin, Dean and Martin 2011, Gabaix and Laibson 2006). Those are represented in a compact way by the inattention parameters m_y, m_r and $\bar{m}$. This paper highlights another potential effect: a "slope" of inattention, whereby agents perceive more dimly things that are further in the future. Gabaix and Laibson (2016) argue that a large fraction of the vast literature on hyperbolic discounting reflects a closely related form of cognitive discounting. Here, and elsewhere, this paper gives functional forms and predictions that can be estimated in future research.

The model could be greatly generalized – for instance, one could generalize $\mathbb{E}_t^{BR}[\boldsymbol{X}_{t+1}] = \bar{m}\boldsymbol{\Gamma}\boldsymbol{X}_t$ to $\mathbb{E}_t^{BR}[\boldsymbol{X}_{t+1}] = diag(\bar{m}_i)\boldsymbol{\Gamma}\boldsymbol{X}_t$, where $diag(\bar{m}_i)$ would be a diagonal matrix of components. In this paper, I try to find a happy medium between psychology and tractability, and this one-dimensional "slope of attention" parametrization proved useful.

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.

Consumption. I derive the consumption of the behavioral agent. To signify "up to the second order term," I use the notation $O\left(\|x\|^2\right)$, where $\|x\|^2 := \mathbb{E}[\hat{y}_\tau^2]/\bar{y}^2 + \mathbb{E}[\hat{r}_\tau^2]/\bar{r}^2$ (the constants $\bar{y}, \bar{r}$ are just here to ensure units are valid).

Proposition 2.1 (Behavioral consumption function) *In this behavioral model, consumption is:* $c_t =$

$c_t^d + \hat{c}_t$, with $c_t^d = \frac{\bar{r}}{R}k_t + \bar{y}$ and

$$\hat{c}_t = \mathbb{E}_t\left[\sum_{\tau\geq t}\frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(b_r\left(k_t\right)m_r\hat{r}_\tau + b_y m_y \hat{y}_\tau\right)\right] + O\left(\|x\|^2\right). \tag{12}$$

$$b_r\left(k_t\right) := \frac{\frac{\bar{r}}{R}k_t - \frac{1}{\gamma}c^d}{R^2}, \qquad b_y := \frac{\bar{r}}{R} \tag{13}$$

Labor supply satisfies: $N_t^\phi = \omega_t c_t^{-\gamma}$, i.e. linearizing

$$\frac{\hat{N}_t}{\bar{N}} = \frac{1}{\phi}\frac{\hat{\omega}_t}{\bar{\omega}} - \frac{\gamma}{\phi}\frac{\hat{c}_t}{c_t^d} + O\left(\|x\|^2\right). \tag{14}$$

The policy of the rational agent is a particular case, setting m_r, m_y and $\bar{m}$ to 1.

The proof is as in Gabaix (2016, Proposition 4.3), which supplies many details about the "boundedly rational dynamic programming" foundations of this equation. The sketch is as follows.[12] First, one derives Proposition 2.1 for the rational agent, i.e. for $m_r = m_y = \bar{m} = 1$. This involves, for instance, a Taylor expansion of the value function. That gives $\hat{c}_t = \mathbb{E}_t\left[\sum_{\tau\geq t}\frac{1}{R^{\tau-t}}\left(b_r\left(k_t\right)\hat{r}_\tau + b_y\hat{y}_\tau\right)\right]$. Next, we use (10), which gives (12).

In (12), consumption reacts to future interest rates and income changes according to the usual income and substitution effects (multiplied by $\frac{1}{\gamma}$). [13,14] Future income is indeed dampened, by a factor $\bar{m}^{\tau-t}$ at horizon $\tau - t$, as in (10).

Note that this agent is "globally patient" for steady-state variables. For instance, his marginal propensity to consume wealth is $\frac{\bar{r}}{R}$, like the rational agent. However, he is myopic to small macroeconomic disturbances in the economy.

This is the microeconomic policy of the representative agent. I next solve for the general equilibrium consequences of this policy.

To simplify notations, I now call $r = R - 1 = \bar{r}$ the steady-state real interest rate.

[12] Auclert (2015) and Woodford (2013) derive related Taylor expansions for the rational model. Section 12.2 contains more details.

[13] Note that here $\hat{y}_\tau$ is the change in total income, which includes changes from endogenous (current and future) labor supply.

[14] Here the labor supply comes from the first order condition. One could develop a more general version $\frac{\hat{N}_t}{\bar{N}} = \frac{\frac{\hat{\omega}_t}{\bar{\omega}} - \gamma m_c^N \frac{\hat{c}_t}{c_t^d}}{\phi}$, where $m_c^N \in [0,1]$ is attention to consumption when choosing labor supply. When $m_c^N = 0$, wealth effects are eliminated. Then, we have a behavioral microfoundation for the labor supply coming traditionally from the Greenwood, Hercowitz and Huffman (1988) preferences, $u\left(C - \frac{N^{1+\phi}}{1+\phi}\right)$.

2.2 Behavioral IS curve: First Without Fiscal Policy

I start with a behavioral New Keynesian IS curve, and consider the case without fiscal policy. The derivation is instructive, and very simple. In the simplest model (this generalizes), the output gap is $x_t = \frac{\hat{c}_t}{c^d}$. Proposition 2.1 gives:

$$x_t = \frac{\hat{c}_t}{c^d} = \frac{1}{c^d}\mathbb{E}_t\left[\sum_{\tau\geq t}\frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(b_y m_y \hat{y}_\tau + b_r\left(k_t\right) m_r \hat{r}_\tau\right)\right]. \tag{15}$$

Now, since there is no capital in the NK model, we have $\hat{y}_\tau = \hat{c}_\tau$: income is equal to aggregate demand. Conceptually, $\hat{c}_\tau$ is the consumption of the other agents in the economy. Hence, using $x_\tau = \frac{\hat{y}_\tau}{c^d}$, we have that with $b_y = \frac{r}{R}$ and $\tilde{b}_r := \frac{b_r(k_t)_{|k_t=0} m_r}{c^d} = -\frac{\frac{1}{\gamma}m_r}{R^2}$, (15) becomes:

$$x_t = \mathbb{E}_t\left[\sum_{\tau\geq t}\frac{\bar{m}^{\tau-t}}{R^{\tau-t}}\left(b_y m_y x_\tau + \tilde{b}_r\hat{r}_\tau\right)\right]. \tag{16}$$

Taking out the first term yields:

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

Given that (16), applied to $t+1$, yields $x_{t+1} = \mathbb{E}_t\left[\sum_{\tau\geq t+1}\frac{\bar{m}^{\tau-t-1}}{R^{\tau-t-1}}\left(b_y m_y x_\tau + \tilde{b}_r\hat{r}_\tau\right)\right]$, we have:

$$x_t = \frac{r}{R}m_y x_t + \tilde{b}_r\hat{r}_t + \frac{\bar{m}}{R}\mathbb{E}_t\left[x_{t+1}\right].$$

Multiplying by R and gathering the x_t terms, we have:

$$x_t = \frac{\bar{m}\mathbb{E}_t\left[x_{t+1}\right] + R\tilde{b}_r\hat{r}_t}{R - r m_y}.$$

This suggests defining $M := \frac{\bar{m}}{R-rm_y}$ and $\sigma := \frac{-R\tilde{b}_r}{R-rm_y} = \frac{m_r/\gamma}{R(R-rm_y)}$. We obtain:

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right] - \sigma\hat{r}_t. \tag{17}$$

which is the "discounted IS curve," with a discount M. The next proposition records the result.

The innovation in the interest rate is written in terms of the nominal rate, i_t, and inflation, π_t:

$$\hat{r}_t := i_t - \mathbb{E}_t [\pi_{t+1}] - r_t^n. \tag{18}$$

Proposition 2.2 (Discounted Euler equation) *In equilibrium, the output gap x_t follows:*

$$x_t = M\mathbb{E}_t [x_{t+1}] - \sigma (i_t - \mathbb{E}_t [\pi_{t+1}] - r_t^n), \tag{19}$$

where $M := \frac{\bar{m}}{R - r m_y} \in [0,1]$ is a modified attention parameter, and $\sigma := \frac{m_r}{\gamma R(R - r m_y)}$. In the rational model, $M = 1$.

Any kind of inattention (to aggregate variables via m_y, cognitive discounting via $\bar{m}$) creates $M < 1$. When the inattention to macro variables is the only force ($\bar{m} = 1$), then $M \in \left[\frac{1}{R}, 1\right]$. Hence, the cognitive discounting gives a potentially powerful quantitative boost.[15]

The behavioral NK IS curve (29) implies:

$$x_t = -\sigma \sum_{\tau \geq t} M^{\tau - t} \mathbb{E}_t [\hat{r}_\tau] \tag{20}$$

i.e. it is the discounted value of future interest rates that matters, rather than the undiscounted sum. This will be important when we study forward guidance below.

Understanding discounting in rational and behavioral models. It is worth pondering where the discounting comes from in (20). 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?

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[16]

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

where $\alpha := \frac{c^d}{\gamma R^2}$. Hence, there is discounting by $\frac{1}{R^\tau}$. However, in general equilibrium, the impact is (see (20) with $M = 1$)

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

[15] 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 (2015)), 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).

[16] See equation (12). I take the case without capital.

so that there is no discounting by $\frac{1}{R^\tau}$. 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 $-\alpha\hat{r}_\tau$. This means that the total impact is larger than the direct effect by a factor

$$\frac{\hat{c}_0^{\mathrm{GE}}}{\hat{c}_0^{\mathrm{direct}}} = R^\tau.$$

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:

$$\text{Behavioral agent: } \hat{c}_0^{\mathrm{direct}} = -\alpha m_r \bar{m}^\tau \frac{\hat{r}_\tau}{R^\tau},$$

which comes from (12). Next, the agent is not fully attentive to indirect effects (including general equilibrium) of future polices. This results in the total effect in (20):

$$\text{Behavioral agent: } \hat{c}_0^{\mathrm{GE}} = -\alpha m_r M^\tau \hat{r}_\tau,$$

with $M = \frac{\bar{m}}{R - r m_y}$. So the multiplier for the general equilibrium effect is:

$$\frac{\hat{c}_0^{\mathrm{GE}}}{\hat{c}_0^{\mathrm{direct}}} = \left(\frac{R}{R - r m_y}\right)^\tau \in [1, R^\tau]$$

and is smaller than the multiplier R^τ in economies with common knowledge of rationality. As m_y becomes 0, the multiplier goes to 1: distant changes in interest rates will be very ineffective if agents are extremely myopic.

2.3 Phillips Curve with Behavioral Firms

Next, I explore what happens if firms do not fully pay attention to future macro variables either.

Firms are the classic Dixit-Stiglitz firms with Calvo pricing frictions. As usual in this literature, I assume a government subsidy to production financed by lump-sum taxes, so that at the steady state the economy is not distorted, and the steady state markup is 0.

I assume that firms are partially myopic to the value of future markups. To do so, I adapt the classic derivation (e.g. chapter 3 of Galí (2015)) for behavioral agents. Firms can reset their prices with probability $1-\theta$ each period. The general price level is p_t, and a firm that resets its price at t sets a price p_t^* according to:

$$p_t^* - p_t = (1-\beta\theta) \sum_{k \geqslant 0}^{\infty} (\beta\theta)^k \, \mathbb{E}_t^{BR} \left[\pi_{t+k} + ... + \pi_{t+1} - \mu_{t+k} \right] \tag{21}$$

i.e. the price is equal to the present value of future marginal costs, which are the future price level minus the markup, μ_{t+k}.

I assume

$$\mathbb{E}_t^{BR} \left[\pi_{t+k} \right] = m_\pi^f \bar{m}^k \mathbb{E}_t \left[\pi_{t+k} - \pi_t^d \right] \tag{22}$$

$$\mathbb{E}_t^{BR} \left[\mu_{t+k} \right] = m_x^f \bar{m}^k \mathbb{E}_t \left[\mu_{t+k} \right] \tag{23}$$

As before, parameters m_π^f and $\bar{m}$ are the "level" and "slope" of the cognitive discounting of attention by firms (which is $m_\pi^f \bar{m}^k$ at horizon k).[17] For generality, I allow for a different intercept of attention to inflation (m_π^f) and the output gap (m_x^f). Otherwise, the setup is as in Galí. When the $m's$ are equal to 1, we have the traditional NK framework.

Tracing out the implications of (21), the macro outcome is as follows.

Proposition 2.3 (Phillips curve with behavioral firms) *When firms are partially inattentive to future macro conditions, the Phillips curve becomes:*

$$\pi_t = \beta M^f \mathbb{E}_t \left[\pi_{t+1} \right] + \kappa x_t \tag{24}$$

with the attention coefficient $M^f := \bar{m} \left(\theta + m_\pi^f (1-\theta) \right)$ *and* $\kappa = \bar{\kappa} m_x^f$*, where* $\bar{\kappa}$ *(given in (86)) is*

[17] I could write $\bar{m}^f$ in (21) rather than $\bar{m}$, at it is the cognitive discounting of firms rather than workers. I didn't do that do avoid multiplying notations, but the meaning should be clear. In particular, in the expression for M^f in (32), it is the $\bar{m}$ of firms that matters, not that of consumers.

independent of attention. Firms are more forward-looking in their pricing (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 (m^f,$\bar{m}$ are higher). When $m_\pi^f = m_x^f = \bar{m} = 1$ (traditional firms), we recover the usual model, and $M^f = 1$.

In the traditional model, the coefficient on future inflation in (24) is exactly β and, miraculously, does not depend on the adjustment rate of prices θ. In the behavioral model (with $m_\pi^f < 1$), in contrast, the coefficient (βM^f) is higher when prices are stickier for longer (higher θ).

Firms can be fully attentive to all idiosyncratic terms (something that would be easy to include), e.g. the idiosyncratic part of their productivity. 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 curve does not change.

The behavioral elements simply change β into βM^f, where $M^f \le 1$ is an attention parameter. Empirically, this "extra discounting" (replacing β by βM^f in (24)) seems warranted, as we shall see in the next section.

Let me reiterate that firms are still fully rational for steady state variables (e.g., in the steady state they 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.

2.4 Extension: Behavioral IS Curve with Fiscal Policy

In this subsection I generalize the above IS curve to the case of an active fiscal policy. The reader is encouraged to skip this section at the first reading.

In this paper, government consumption is always 0. Fiscal policy means cash transfers from the government to agents and lump-sum taxes. 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. It evolves as $B_{t+1} = \frac{1+i_t}{1+\pi_{t+1}} (B_t + \mathcal{T}_t)$ where $\mathcal{T}_t$ is the lump-sum transfer given by the government to the agent (so that $-\mathcal{T}_t$ is a tax), and $\frac{1+i_t}{1+\pi_{t+1}}$ is the realized gross return on bonds.[18] Taking a Taylor expansion,[19] debt evolves as:

$$B_{t+1} = R (B_t + \mathcal{T}_t) .$$

[18]The debt is short-term. Debt maturity choice is well beyond the scope of this paper.

[19]That is, I formally consider the case of "small" debts and deficits, neglecting the variations of the real rate (i.e. second-order terms $O\left(\left|\frac{1+i_t}{1+\pi_{t+1}} - R\right| (|B_t| + |d_t|)\right)$).

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:[20]

$$B_{t+1} = B_t + R d_t.$$

Iterating 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:

$$\mathcal{T}_\tau = -\frac{r}{R} B_t + \left(d_\tau - r \sum_{u=t}^{\tau-1} d_u \right). \tag{25}$$

Equation (25) is the objective law of motion of the transfer. The general formalism gives the following behavioral version of that equation, as perceived by the agent:[21]

$$\mathbb{E}_t^{BR} \left[\mathcal{T}_\tau \right] = -\frac{r}{R} B_t + m_y \bar{m}^{\tau-t} \left(d_\tau - r \sum_{u=t}^{\tau-1} d_u \right). \tag{26}$$

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 $\mathbb{E}_t^{BR} \left[\mathcal{T}_\tau \right] = -\frac{r}{R} B_t$. However, he sees future deficits dimly. This is captured by the term $m_y \bar{m}^{\tau-t}$.

Consumption satisfies, again from Proposition 2.1:

$$x_t = \mathbb{E}_t^{BR} \left[\sum_{\tau \geq t} \frac{1}{R^{\tau-t}} \left(b_y \left(x_\tau + \mathcal{T}_\tau \right) + b_r \left(k_t \right) \hat{r}_\tau \right) \right] + \frac{r}{R} k_t$$

where $\mathbb{E}_t^{BR}$ is the expectation under the subjective model and in equilibrium $k_t = B_t$.

Calculations in the Appendix (section 8) give the following modifications of the IS curve. Note that here we only have deficits, not government consumption.

Proposition 2.4 (Discounted Euler equation with sensitivity to budget deficits) *We have the following IS curve reflecting the impact of both fiscal and monetary policy:*

$$x_t = M \mathbb{E}_t \left[x_{t+1} \right] + b_d d_t - \sigma \left(i_t - \mathbb{E}_t \pi_{t+1} - r_t^n \right) \tag{27}$$

[20] Indeed, $B_{t+1} = R \left(B_t - \frac{r}{R} B_t + d_t \right) = B_t + R d_t$.
[21] See the derivation of Proposition 2.4 for details.

where d_t is the budget deficit and

$$b_d = \frac{rm_y}{R - m_y r} \frac{R(1-\bar{m})}{R-\bar{m}} \tag{28}$$

is the sensitivity to deficits. When agents are rational, $b_d = 0$, but with behavioral agents, $b_d > 0$.

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

Here I assume a representative agent. This analysis complements analyses that assume heterogeneous agents to model non-Ricardian agents, in particular rule-of-thumb agents à la Campbell-Mankiw (1989), Galí, López-Salido and Vallés (2007), Mankiw (2000), Bilbiie (2008), Mankiw and Weinzierl (2011) and Woodford (2013).[22,23] When dealing with complex situations, a representative agent is often simpler. In particular, it allows us to value assets unambiguously.[24]

2.5 Synthesis: Behavioral New Keynesian Model

I now gather the above results.

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

$$x_t = M\mathbb{E}_t[x_{t+1}] + b_d d_t - \sigma(i_t - \mathbb{E}_t\pi_{t+1} - r_t^n) \text{ (IS curve)} \tag{29}$$

$$\pi_t = \beta M^f \mathbb{E}_t[\pi_{t+1}] + \kappa x_t \text{ (Phillips curve)} \tag{30}$$

where $M, M^f \in [0,1]$ are the attention of consumers and firms, respectively, to macroeconomic

[22] Galí, López-Salido and Vallés (2007) model is richer and more complex, as it features heterogeneous agents. Omitting the monetary policy terms, instead of $x_t = \mathbb{E}_t\left[\sum_{\tau \geq t} \frac{M^{\tau-t}}{R^{\tau-t}} b_d d_\tau\right]$ (see (27)), they generate $x_t = \Theta_n n_t - \Theta_\tau t_t^r$, where t_t^r are the rational agents. Hence, one key difference is that in the present model, future deficits matter as well, whereas in their model, they do not.

[23] 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. Another way to have non-Ricardian agents is via rational credit constraints, as in Kaplan, Moll and Violante (2016). The analysis is then rich and complex.

[24] With heterogenous agents and incomplete markets, there is no agreed-upon way to price assets: it is unclear "whose pricing kernel" one must take.

outcomes, and $b_d \geq 0$ is the impact of deficits:

$$M = \frac{\bar{m}}{R - rm_y}, \quad b^d = \frac{rm_y}{R - m_y r}\frac{R(1-\bar{m})}{R - \bar{m}}, \tag{31}$$

$$M^f = \bar{m}\left(\theta + m_\pi^f (1-\theta)\right), \quad \kappa = \bar{\kappa} m_x^f. \tag{32}$$

In the traditional model, $\bar{m} = m_y = m_r = m_\pi^f = m_x^f = 1$, so that $M = M^f = 1$ and $b_d = 0$. In addition, $\sigma := \frac{m_r}{\gamma R(R - rm_y)}$, and $\bar{\kappa}$ (given in (86)) is independent of attention.

The reader may be bewildered by a model with five behavioral parameters. Fortunately, only one is very crucial—the cognitive discounting factor $\bar{m}$. The other four parameters $(m_y, m_r, m_x^f, m_\pi^f)$ are not essential, and could be set to 1 (the rational value) in most cases. Still, I keep them here for two reasons: conceptually, I found it instructive to see where the intercept, rather than the slope of attention, matters. Also given those various "intercepts of attention" are conceptually natural, they are likely to be empirically relevant as well when future studies measure attention.[25]

Empirical Evidence on the Model's Deviations from Pure Rationality. The empirical evidence, we will now see, appears to support the main deviations of the model from pure rationality.

In the Phillips curve, firms do not appear to be fully forward looking: $M^f < 1$. Empirically, the Phillips curve is not very forward looking. For instance, Galí and Gertler (1999) find that we need $\beta M^f \simeq 0.84$ at the quarterly frequency; given that $\beta \simeq 0.99$, that leads to an attention parameter of $M^f \simeq 0.85$. If $m_\pi^f = 1$, this corresponds to $\bar{m} = 0.85$. This is the value I will take in the numerical examples, fully detailed in Section 9.8.

In the Euler equation consumers do not appear to be fully forward looking: $M < 1$. The literature on the forward guidance puzzle concludes, plausibly I think, that $M < 1$.

Ricardian equivalence does not fully hold. There is much debate about Ricardian equivalence. The provisional median opinion is that it only partly holds. For instance, the literature on tax rebates (Johnson, Parker and Souleles 2006) appears to support $b^d > 0$.

Indeed, all three facts come out naturally from a model with cognitive discounting $\bar{m} < 1$.

[25] In continuous time, we write $M = 1 - \xi \Delta t$ and $\beta M^f = 1 - \rho \Delta t$. In the small time limit ($\Delta t \to 0$), $\xi \geq 0$ is the cognitive discounting parameter due to myopia in the continuous time model, while ρ is the discount rate inclusive of firm's myopia. The model (29) becomes, in the continuous time version:

$$\dot{x}_t = \xi x_t - b_d d_t + \sigma (i_t - r_t - \pi_t) \tag{33}$$

$$\dot{\pi}_t = \rho \pi_t - \kappa x_t. \tag{34}$$

When $\xi = b_d = 0$, we recover Werning (2012)'s formulation, which has rational agents.

Related work Models with hand-to-mouth agents generate $M = 1$.[26] McKay, Nakamura and Steinsson (2016) provide a microfoundation for an approximate version of this IS curve (29) based on heterogeneous rational agents with limited risk sharing (see also Campbell et al. 2016), without an analysis of deficits d_t. Werning (2015)'s analysis yields a modified Euler equation with rational heterogeneous agents, which often yields $M > 1$. Del Negro, Giannoni and Patterson (2015) and Eggertsson and Mehrotra (2015) microfoundations with from the finite-lives of agents. Finite lives severely limits how myopic agents can be (e.g. predicts an M very close ot 1), given that life expectancies are quite high.[27]

Caballero and Farhi (2015) offer a different explanation of the forward guidance puzzle in a model with a shortage of safe assets and endogenous risk premia and (see also Caballero, Farhi and Gourinchas 2015). Relatedly, Fisher (2015) derives a discounted Euler equation with a safe asset premium: but the effect is very small, e.g. the coefficient $1 - M$ is very close to 0 – close to the empirical "safety premium", so at most $M = 0.99$.

My take, in contrast, is behavioral: the reason that forward guidance does not work well is that it is in some sense "too subtle" for the agents. This allows to microfound very simply not only a discounted Euler equation, but also a discounted Phillips curve, with $M^f < 1$, and in both cases with M potentially much below 1. In independent work, Garcia-Schmidt and Woodford (2015) offer another, distinct, behavioral take on the IS curve.

We now study several consequences of these modifications for the forward guidance puzzle.

3 Consequences of this Behavioral Model

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

[26] 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 $M = 1$.

[27] The Mankiw-Reis (2002) model changes the Phillips curve, which helps for some paradoxes (Kiley 2016). But as it keeps the same IS curve as the traditional model, it still requires $\phi_\pi > 1$ for determinacy – unlike the present behavioral model. A synthesis of Mankiw-Reis and the present model would be useful.

strong impacts for the economy at the ZLB.

I assume that the central bank follows a Taylor rule of the type:

$$i_t = \phi_\pi \pi_t + \phi_x x_t + j_t \tag{35}$$

where j_t is typically just a constant.[28] It is useful to define the state vector[29]

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

and $a_t := j_t - r_t^n$ (as in "*a*ction") the baseline tighteness of monetary policy (if the government pursues the first best, $a_t = 0$). For simplicity, I assume an inactive fiscal policy, $d_t = 0$.[30] Calculations yield:[31]

$$\boldsymbol{z}_t = \boldsymbol{A}\mathbb{E}_t [\boldsymbol{z}_{t+1}] + \boldsymbol{b} a_t \tag{36}$$

with

$$\begin{aligned} \boldsymbol{A} &= \frac{1}{1+\sigma(\phi_x + \kappa\phi_\pi)} \begin{pmatrix} M & \sigma\left(1-\beta^f \phi_\pi\right) \\ \kappa M & \beta^f (1+\sigma\phi_x) + \kappa\sigma \end{pmatrix}, \\ \boldsymbol{b} &= \frac{-\sigma}{1+\sigma(\phi_x+\kappa\phi_\pi)} (1,\kappa)'. \end{aligned} \tag{37}$$

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

Proposition 3.1 (Equilibrium determinacy with behavioral agents) *There is a determinate equi-*

[28] The reader will want to keep in mind the case of a constant $j_t = \bar{j}$. More 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] I call $\boldsymbol{z}_t$ the "state vector" with some mild abuse of language. It's an outcome of the deeper state vector $\boldsymbol{X}_t$. In stability analysis, it's conventionally to call it a state vector.

[30] Given a rule for fiscal policy, the sufficient statistic is the behavior of the "monetary and fiscal policy mix" $i_t - \frac{b_d d_t}{\sigma}$. For instance, suppose that: $i_t - \frac{b_d d_t}{\sigma} := \phi_\pi \pi_t + \phi_x x_t + j_t$, with some (unimportant) decomposition between i_t and d_t. The analysis is then the same. More general analyses might add the total debt D_t as a state variable in the rule for d_t.

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

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

librium (all of $\boldsymbol{A}$'s eigenvalues are less than 1 in modulus) if and only if:

$$\phi_\pi + \frac{\left(1-\beta^f\right)}{\kappa}\phi_x + \frac{\left(1-\beta^f\right)(1-M)}{\kappa\sigma} > 1. \tag{38}$$

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[32]

$$\text{Strong enough bounded rationality condition: } \frac{\left(1-\beta M^f\right)(1-M)}{\kappa\sigma} > 1. \tag{39}$$

Condition (39) does not hold in the traditional model, where $M = 1$. The condition basically means that agents are boundedly rational enough, that is M is sufficiently less than 1 – and the firm-level frictions cognitive and pricing frictions ($\kappa = \bar{\kappa} m_x^f$) are large enough.[33]

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

Condition (39) 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 (39) is violated in the traditional model).

This absence of multiple equilibria is important. Indeed, take a central bank following a fixed interest path – for instance in a period of a prolonged ZLB. Then, in the traditional model, there is always a continuum of (bounded) equilibria, technically, because 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. In contrast, in

[32]In continuous time, criterion (38) is: $\phi_\pi + \frac{\rho}{\kappa}\phi_x + \frac{\rho\xi}{\kappa\sigma} > 1$, so that condition (39) is: $\frac{\rho\xi}{\kappa\sigma} > 1$.

[33]Recall also that $\kappa = \bar{\kappa} m_x^f$. So, greater bounded rationality by firms (lower m_x^f) helps achieving unicity. As the frequency of price changes becomes infinite, $\kappa \to 0$ (see equation (86)). 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 Kocherlakota (2016)'s finding that we need enough price stickiness to have sensible predictions in long-horizon models.

[34]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.

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

this behavioral model, we can simply write:

$$\boldsymbol{z}_t = \mathbb{E}_t \left[\sum_{\tau \geq t} \boldsymbol{A}^{\tau - t} \boldsymbol{b} a_\tau \right]. \tag{40}$$

Cochrane (2015) made the point that we'd expect the economy such as Japan to be quite volatile, if the ZLB is expected to last forever: conceivably, the economy could jump from one equilibrum to the next at each period. Indeed, let us add a bit to his point, and observe that the economy would be also very volatile even if the ZLB lasts for a finite but long amount of time. Indeed, suppose that the ZLB is expected to last for T periods. Call $\boldsymbol{A}_{ZLB}$ the value of matrix $\boldsymbol{A}$ in (37) when $\phi_\pi = \phi_x = j = 0$ in the Taylor rule. Then, the system (36) is, at the ZLB ($t \leq T$): $\boldsymbol{z}_t = \mathbb{E}_t \boldsymbol{A}_{ZLB} \boldsymbol{z}_{t+1} + \underline{\boldsymbol{b}}$ with $\underline{\boldsymbol{b}} := (1, \kappa)\, \sigma \underline{r}$. Hence, iterating forward, we have:

$$\boldsymbol{z}_0 (T) \quad = \quad \left(\boldsymbol{I} + \boldsymbol{A}_{ZLB} + ... + \boldsymbol{A}_{ZLB}^{T-1} \right) \underline{\boldsymbol{b}} + \boldsymbol{A}_{ZLB}^T \mathbb{E}_0 \left[\boldsymbol{z}_T \right] \tag{41}$$

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 today to $\mathbb{E}_0 [\boldsymbol{z}_T]$, i.e. to the expectations after the ZLB, have unboundedly large impact today ($\lim_{T \to \infty} \| \boldsymbol{A}_{ZLB}^T \| = \infty$) . Hence, we would expect the economy to be very unstable 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 $\underline{r} < 0 < \bar{r}$. I assume that for $t > T$, the central bank implements $x_t = \pi_t = 0$ by setting $i_t = \bar{r}$. At time $t < T$, I suppose that the CB is at the ZLB, so that $i_t = 0$.

Proposition 3.2 *In the traditional rational case* ($M = 1$)*, we obtain an unboundedly intense recession as the length of the ZLB increases:* $\lim_{t \to -\infty} x_t = -\infty$*. This also holds when myopia is mild, i.e. (39) fails.*

However, suppose cognitive myopia is strong enough, i.e. (39) holds. Then, we obtain a bound-

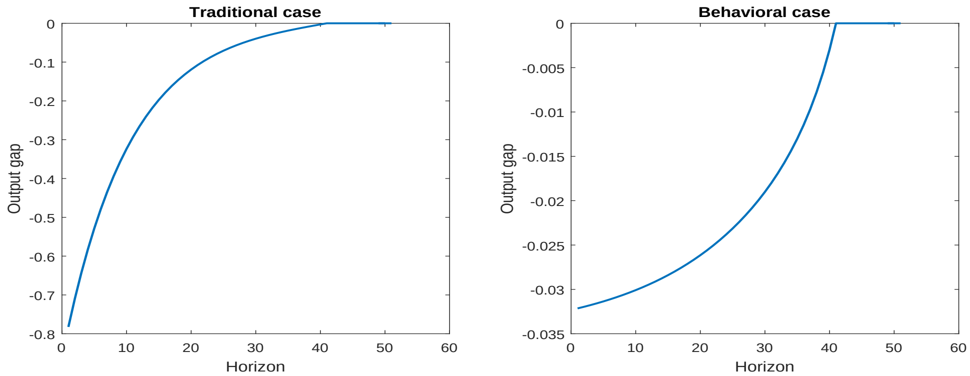

Figure 1: This Figure shows the output gap x_t. The economy is at the Zero Lower Bound during times 0 to $T = 40$ quarters. The left panel is the traditional New Keynesian model, the right panel the present behavioral model. Parameters are the same in both models, except that attention is $M = 0.85$ in the present behavioral model, and $M = 1$ in the traditional model. Units: the output gap at time 0 is -78 in the traditional case, and -3.2 in this behavioral case. Time units are quarterly.

edly intense recession: $\lim_{t\to-\infty} x_t = \frac{\sigma(1-\beta^f)}{(1-M)(1-\beta^f)-\kappa\sigma}\underline{r} < 0$.

We see how impactful myopia can be. We see that 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.

Figure 1 shows the dynamics.[36] The left panel shows the traditional model, the right one the behavioral model. The parameters are the same in both models, except that attention is lower (set to an annualized rate of 85) in the behavioral model (against its value $M = 1$ in the traditional 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, while 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 1945-46 or Rwanda.

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.

[36] The numerical values are detailed in Section 9.8.

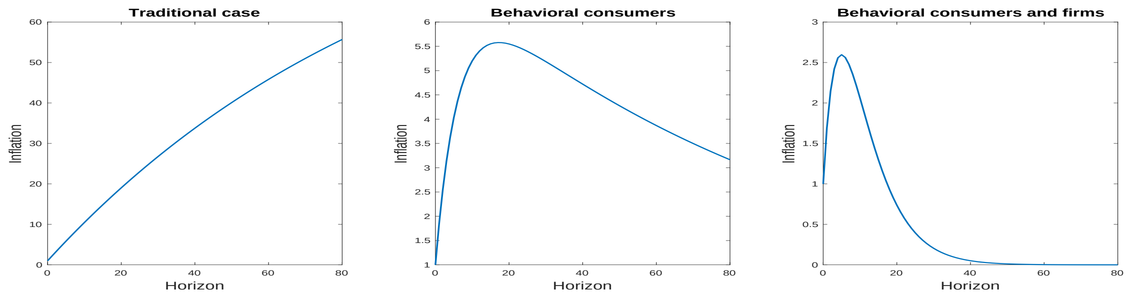

Figure 2: This Figure shows the response of current inflation to forward guidance about a one-period interest rate cut in T quarters, compared to an immediate rate change of the same magnitude. Left panel: traditional New Keynesian model. Middle panel: model with behavioral consumers and rational firms. Right panel: model with behavioral consumers and firms. Parameters are the same in both models, except that attention is $M = M^f = 0.85$ in the behavioral model, and $M = 1$ in the traditional model.

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 tax cut. What is the impact on today's inflation? This is the thought experiment analyzed by McKay, Nakamura and Steinsson (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 way the policy, the bigger the impact today – this 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 distant policy changes have vanishingly small effects with behavioral agents – but they have the biggest effect with rational agents. Bounded rationality of both firms and consumers is useful for the effect.

Formally, we have $x_t = Mx_{t+1} - \sigma \hat{r}_t$, 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 inflation 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 $M < 1$).[37]

[37]When attention is endogenous, the analysis could become more subtle. Indeed, if other agents are more attentive

I close here the discussion of the empirical advantages of having the strong myopia conditon (39). Cochrane (2016a) discusses more such advantages – and concludes that we need either a behavioral New Keynesian model like this one, or a very different theory that he exposes in Cochrane (2016b), based on the fiscal theory of the price level. I next turn to optimal policy in this behavioral model.

4 Optimal Monetary and Fiscal Policy

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

Optimal policy needs a welfare criterion. Welfare here is 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 behavioral economics, which views behavioral agents as using heuristics, but experience utility from consumption and leisure like rational agents.[38] I express $\widetilde{W} = W^* + W$, where W^* is first best welfare, and W is the deviation from the first best. The following lemma (whose proof in the online appendix, Section 10) derives it.

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

$$W = -K\mathbb{E}_0 \sum_{t=0}^{\infty} \frac{1}{2}\beta^t \left(\pi_t^2 + \vartheta x_t^2\right) + W_- \tag{42}$$

where $\vartheta = \frac{\bar{\kappa}}{\varepsilon} = \frac{\kappa}{m_x^f \varepsilon}$, $K = u_c c\left(\gamma + \phi\right)\frac{\epsilon}{\bar{\kappa}}$, *and* W_- *is a constant (explicited in (106)),* $\bar{\kappa}$ *is independent of bounded rationality,* $\kappa = m_x^f \bar{\kappa}$ *is Phillips curve coefficient,* ε *is the elasticity of demand, and* $m_x^f \in (0, 1]$ *is firms' attention to the output determinant of the markup.*

Hence, the welfare losses are the same as in the rational model, when expressed in terms of deep parameters (including $\bar{\kappa}$). However, when expressed in terms of the Phillips curve coefficient κ, the relative weight on the output gap (ϑ) is higher when firms are more behavioral (when m^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 lemma helps alleviate.

to the forward Fed announcement, 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, m_r. I do not pursue that here.

[38] In particular I use the objective (not subjective) expectations. Also, I do not count thinking costs in the welfare. 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 (2015), we are in the "no attention in welfare" case.

4.2 Optimal policy: Response to Changes in the Natural Interest Rate

Suppose that there is a productivity or discount factor shocks (the latter are not explicitly in the basic model, but can be introduced straightforwardly). This changes the natural real interest rate, r_t^n.[39]

To find the policy ensuring the first best (i.e. 0 output gap and inflation), we inspect the two equations of this behavioral model (equations (29)-(30)). This reveals that the first best is achieved if and only if: $i_t - \frac{b_d}{\sigma} d_t = r_t^n$. We record this.

Proposition 4.2 (First best) *When there are shocks to the natural rate of interest, the first best is achieved if and only if:*

$$i_t - \frac{b_d}{\sigma} d_t = r_t^n. \tag{43}$$

This means that if the economy has a lower natural interest rate (hence "needs loosening"), the government can either decrease rates, or increase deficits. Monetary and fiscal policy are substitutes.[40]

4.2.1 When the ZLB doesn't bind: Monetary policy attains the first best

Suppose that the ZLB doesn't bind ($r_t^n \geq 0$). Then, we can turn off fiscal policy ($d_t = 0$). With rational and behavioral agents, the optimal policy is still to set $i_t = r_t^n$, i.e. to make the nominal rate track the natural real rate: [41,42]

$$\text{First best away from the ZLB: } i_t = r_t^n \text{ and zero deficit: } d_t = 0. \tag{44}$$

This is the traditional, optimistic message in monetary policy.[43]

[39] Bounded rationality modulate the way TFP and preference shocks change the natural interest rate (see Section 9.1). In a model with capital, the situation would be more complex.

[40] If there are budget deficits, the central bank must "lean against behavioral biases". 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 ouput gap (in the IS curve (29)) by setting: $i_t = r_t^n + \frac{b_d}{\sigma} d_t$.

[41] As is well understood, to ensure equilibrium determinacy, the central bank embeds this in-sample policy into a more general rule, e.g. sets $i_t = r_t^n + \phi_\pi \pi_t + \phi_x x_t$ with coefficients ϕ_π, ϕ_x sufficiently large (following (38)). On the equilibrium path, $\pi_t = x_t = 0$, so that $i_t = r_t^n$.

[42] If the inflation target was $\bar{\pi}$, 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 actually policy is: $i_t = i_t^* + \phi_x (x_t - x_t^*) + \phi_\pi (\pi_t - \pi_t^*)$ with coefficients ϕ that satisfies the modified Taylor criterion (38).

4.2.2 When the ZLB binds: "Helicopter drops of money" as an optimal cure in the Optimal Mix of Fiscal and Monetary Policy

However, when the natural rate becomes negative (and with low inflation), the optimal nominal interest rate is negative, which is by and large not possible.[44] That is the ZLB. The first best is not achievable in the traditional model. Then, much research has shown that the policy is quite complex then. [45]

However, with behavioral agents, there is an easy first best policy. This first best policy is as follows:[46]

$$\text{First best at the ZLB: } i_t = 0 \text{ and deficit: } d_t = \frac{-\sigma}{b_d} r_t^n \tag{45}$$

Fiscal policy runs deficits, to stimulate demand. By "fiscal policy" I mean transfers (from the government to the agents), and "helicopter drops of money", i.e. checks that the central bank might send (this gives some fiscal authority to the central bank).[47] This is again possible because agents are non Ricardian.

In Appendix 9.2 I analyze a richer situation, and shows that the possibility of future fiscal policy can have ex ante benefits – it makes agents confident about the future, as the know that the government will not run out of tools.

In conclusion, behavioral considerations considerably change policy at the ZLB, and allow to achieve the first best.[48]

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 next 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

[44]Recent events have seen nominal rates slightly below 0%, but it does not seem possible to obtain very low nominal rates, say -5%, for long, because stockpiling cash in a vault is then a viable alternative.

[45]The first best is not achievable, and second best policies are complex, as been analyzed by a large number of authors, e.g. Eggertson and Woodford (2003), Werning (2012) and the survey in Galí (2015, section 5.4).

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

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

[48]Models with rational agents and credit constraints might work similarly (e.g. Kaplan, Moll Violante 2016). However, they will be much harder to analyze, in part because future policy will have a complex, non linear effect savings and the like.

disturbance ν_t to the Phillips curve, which becomes:

$$\pi_t = \beta M^f \mathbb{E}_t [\pi_{t+1}] + \kappa x_t + \nu_t \quad (46)$$

andν_t follows an AR(1), $\nu_t = \rho_\nu \nu_{t-1} + \varepsilon_t^\nu$. 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 Galí (2015, Section 5.2) for microfoundations).[49]

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

4.3.1 Optimal commitment policy

The next proposition states the optimal policy with commitment.

Proposition 4.3 (Optimal policy with commitment: suboptimality of price-level targeting) *The optimal commitment policy entails:*

$$\pi_t = \frac{-\vartheta}{\kappa} \left(x_t - M^f x_{t-1} \right) \quad (47)$$

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

$$p_t = \frac{-\vartheta}{\kappa} \left(x_t + \left(1 - M^f\right) \sum_{\tau=0}^{t-1} x_\tau \right) \quad (48)$$

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). However, with behavioral firms, the price level goes up (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

[49] Analyzing an early version of the present model, Bounader (2016) examined various constrained policies and derived independently some results of this section 4.3, though not the key result on the non-optimality of price-level targeting.

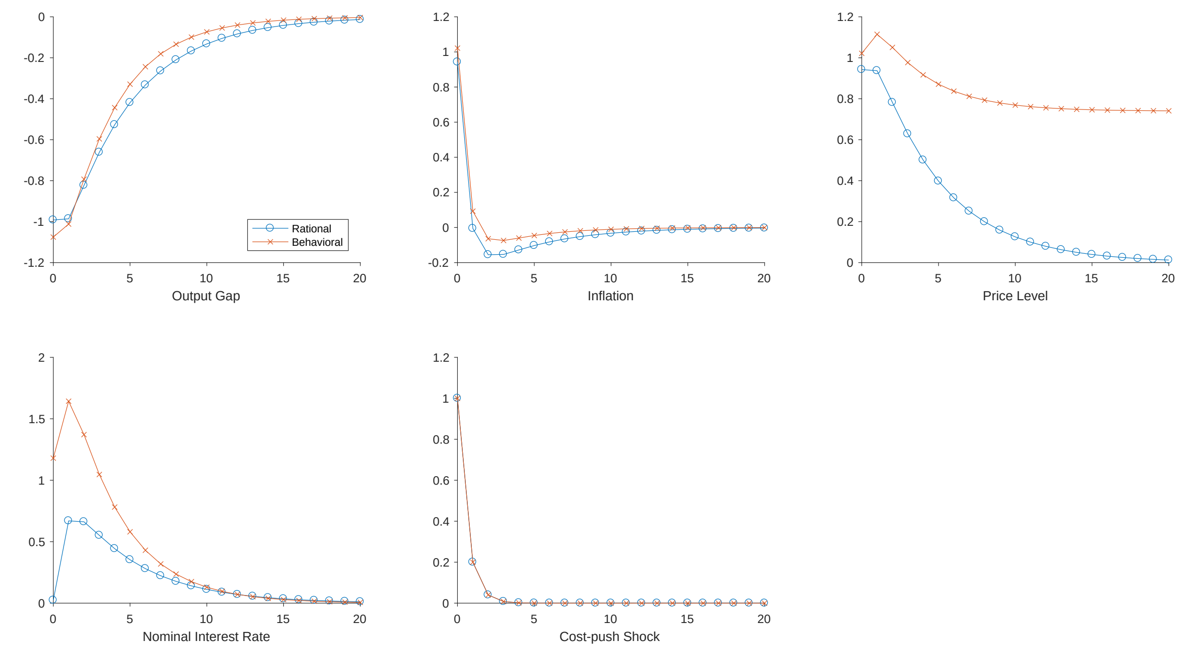

Figure 3: This figure shows optimal interest rate policy in response to a cost-push shock (ν_t), when the central bank follows the optimal commitment strategy. When firms are rational, the optimal strategy entails "price level targeting", i.e. the central bank will engineer a deflation later to come back to the initial price level. This is not the optimum policy with behavioral firms. This illustrates Proposition 4.3.

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.[50] This lack of intuitive justification may be caused by that fact that it's not robust to behavioral deviations, as Proposition 4.3 shows.

Likewise, "nominal GDP targeting" is optimal in the traditional model, but it is suboptimal with behavioral agents.

Other considerations The gains from commitment are lower, as firms don't react much to the future. At the same time, the optimal policy still features "history dependence" (in the terminology of Woodford 2003), even when the cost-push shock has no persistence: see equation (47).

Figure 3 gives some more intuition. Look at the behavioral of the interest rate. The policy 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.

4.3.2 Optimal discretionary policy

Proposition 4.4 (Optimal discretionary policy) *The optimal discretionary policy entails:*

$$\pi_t = \frac{-\vartheta}{\kappa} x_t \tag{49}$$

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

Let us first examine the comparative statics controlling for κ. For transitory shocks ($\rho_\nu = 0$), the optimal policy is independent of the firms' bounded rationality. Future considerations don't matter. However, for persistent shocks, the optimal policy is less aggressive ($\frac{di_t}{d\nu_t}$ is lower) when firms are more behavioral. This is because, with more myopic firms, future cost push shocks do not affect firms' pricing today much, hence the central bank needs to respond less to them.

[50] This is not a particularly intuitive fact, 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 are the same. That identity is broken in the behavioral model. This is analogous to Slutsky symmetry in the rational model: there is no great intuition for its justification in 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.

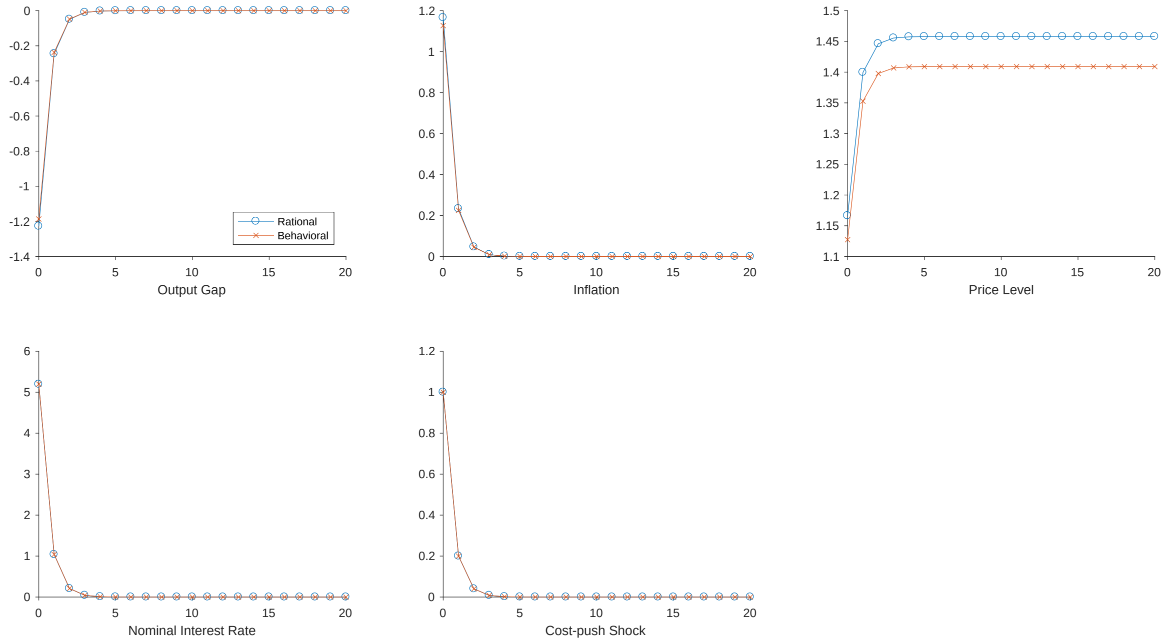

Figure 4: This figure shows optimal interest rate policy in response to a cost-push shock (ν_t), when the central bank follows the optimal discretionary strategy. The behavior is very close, as the central bank does not rely on future commitments for its optimal policy. This illustrates Proposition 4.4.

5 Enriching the Model with Long-Run Changes to Inflation

5.1 Enriched Model

So far all variables came back to a steady state value normalized to 0. This is sufficient for most of the analysis. Here, I extend the analysis to allow for the possibility that long run inflation might change. For good measure, I also extend the model to have backward looking terms, which has proven useful in empirical analyses, as this creates inertia in inflation (e.g. Galí and Gertler 1999). The interaction of backward looking terms for firms and permanent changes will prove fruitful.[51]

I propose to modify (22) into:

$$\mathbb{E}_t^{BR}\left[\pi_{t+k}\right] \quad = \quad \pi_t^d + m_{f,\pi}\bar{m}^k \mathbb{E}_t\left[\pi_{t+k} - \pi_t^d\right] \tag{50}$$

where π_t^d is a "default inflation" which follows (53) below. This means that when a firm predicts future inflation π_{t+k}, it first anchors it on default inflation π_t^d, like in Tversky and Kahneman

[51] Galí and Gertler (1999) present a model with partially backward looking firms: their model has $\eta = 1$, $M = 1$. However, they have $\zeta = 0$, which prevents the stability analysis below, where $\zeta > 0$ is crucial.

(1974)'s anchoring and adjustment. Then, it partially adjusts for the deviation of future inflation from that default value – that's the term $\mathbb{E}_t\left[\pi_{t+k}-\pi_t^d\right]$. The term $m_{f,\pi}\bar{m}^k \leq 1$ means that the adjutment is partial.

Equation (53) means that default inflation is a mix of inflation guidance by the central bank π_t^{CB}, and past inflation. If the central bank keeps says "our long term inflation term is 2%", then $\pi_t^{CB}=2\%$. Firms put a weight $\zeta\in[0,1]$ on the central bank guidance, and $1-\zeta$ on past inflation.[52]

This is the only modification of the baseline model, which is the particular case $\pi_t^d\equiv 0$. Working out the general equilibrium leads to the following.

Proposition 5.1 (Behavioral New Keynesian model – three equation version) *We obtain the following behavioral version of the New Keynesian model, for the behavior of output gap x_t and inflation π_t:*

$$x_t = M\mathbb{E}_t\left[x_{t+1}\right]+b_d d_t-\sigma\left(i_t-\mathbb{E}_t\pi_{t+1}-r_t^n\right) \tag{51}$$

$$\pi_t = \beta^f\mathbb{E}_t\left[\pi_{t+1}\right]+\alpha^f\left(\pi_t^d-\beta\theta\bar{m}\pi_{t+1}^d\right)+\kappa x_t \tag{52}$$

$$\pi_{t+1}^d = (1-\eta)\,\pi_t^d+\eta\left(\zeta\pi_t^{CB}+(1-\zeta)\,\pi_t\right) \tag{53}$$

with all terms as before, and $\alpha^f\geq 0$ is given in (87). The new terms are π_t^d, "default inflation" and π_t^{CB}, the "inflation guidance" by the central bank.

Leading old and new Keynesian models are embedded in the structure (51)–(53), as we shall see. It will be convenient to define

$$\alpha := \alpha^f\left(1-\beta\theta\bar{m}\right)$$

which verifies $\alpha+\beta^f\leq\beta$. Equation (52) shows that if π^d is constant, the impact on π_t is $\alpha\pi^d$.

5.2 Long Run Behavior and Determinacy

Given our system (51)-(53), we ask two questions: does Fisher neutrality (or something close to it) hold? Is the equilibrium determinate with fixed rates (e.g. at the ZLB)? The analysis will reveal a connection between those properties.

[52]Hence, firms has two ways of predicting future inflation: one is via "purely rational expectations", with π_t, another is via default inflation, π_t^d. I view this default inflation as a simpler, easily available source of signal about future inflation. In a noisy signal model (Velkamp 2011), π_t^d would be a precise signal, while the firm would get only a noisy signal about future inflation $\mathbb{E}_t\left[\pi_{t+k}\right]$.

We first make a few observations. Consider the long run value of inflation (π_∞) and the nominal rate (i_∞). Their link is as follows.

Proposition 5.2 (Long run Fisher neutrality) *If long run inflation is higher by $d\pi_\infty$, then the long run nominal rate is higher by di_∞, where:*

$$\frac{di_\infty}{d\pi_\infty} = 1 - \frac{\left(1-\alpha-\beta^f\right)(1-M)}{\kappa\sigma} \tag{54}$$

Proof. Just plug constant values of $\pi_t = \pi_t^d = \pi_t^{CB} = \pi_\infty$, and $x_t = x_\infty$ into (51)-(53), which gives $\frac{i_\infty}{\pi_\infty} = 1 - \frac{\left(1-\alpha-\beta^f\right)(1-M)}{\kappa\sigma}$. In addition, $x_\infty = \frac{\left(1-\alpha-\beta^f\right)}{\kappa}\pi_\infty$. $\square$

Next, I ask: is the equilibrium determinate?[53] The next Proposition generalizes the earlier criterion (Proposition 3.1) to the case with backward looking terms.[54]

Proposition 5.3 (Equilibrium determinacy with behavioral agents – with backward looking terms) *The equilibrium is determinate only if:*

$$\phi_\pi + \frac{\left(1-\alpha(1-\zeta)-\beta^f\right)(1-M+\sigma\phi_x)}{\kappa\sigma} > 1. \tag{55}$$

Now, which properties of real economies should a model reflect? First, in the long run, a steady state rise of in the nominal rate is associated with a rise in inflation: $\frac{di_\infty}{d\pi_\infty} > 0$, something we might call "long run Fisher sign neutrality" (pure Fisher neutrality would be $\frac{di_\infty}{d\pi_\infty} = 1$). Most studies (e.g. Kandel, Ofer and Sarig 1996, Evans 1998) find $\frac{di_\infty}{d\pi_\infty} > 0$ – though typically also they reject pure Fisher neutrality ($\frac{di_\infty}{d\pi_\infty} = 1$), and instead find $\frac{di_\infty}{d\pi_\infty} < 1$, qualitatively as in Proposition 5.2.[55] This means, given (54), that the data wants:

$$\frac{\left(1-\alpha-\beta^f\right)(1-M)}{\kappa\sigma} < 1. \tag{56}$$

Second, in the recent experience in Japan (since the late 1980s) and in Europe and US (since 2010), the interest rate has been stuck at the ZLB, but without strong vagaries of inflation or output.

[53] Mathematically, this is the following. Define $\boldsymbol{z}_t := \left(x_t, \pi_t, \pi_t^d\right)$, taking π_t^{CB} as given, and write the system as $\mathbb{E}_t \boldsymbol{z}_{t+1} = \boldsymbol{B}\boldsymbol{z}_t + a\pi_t^{CB}$, for a matrix $\boldsymbol{B}$. Given that π_t^d is a predetermined variable, we need $\boldsymbol{B}$ to have 2 eigenvalues with modulus greater than 1, and 1 with modulus less than 1.

[54] This proposition states a necessary condition. The necessary and sufficient condition is (55) and a "Routh-Hurwitz auxiliary condition" stated in the online appendix (Proposition 9.3). This condition is much more minor, and almost automatically valid in practice (see the discussion in the online appendix).

[55] The identification problems are very difficult, in part because we deal with fairly long run outcomes on which there are few observations.

Hence, following Cochrane (2015), I hypothesize that another desirable empirical "target" for the model is that the equilibrium is determinate (the economy is stable) even if the monetary policy is stuck at the ZLB forever.[56] In the model, that means that (use (55) in the case $\phi_\pi = \phi_x = 0$):

$$\frac{\left(\alpha\zeta + 1 - \alpha - \beta^f\right)(1 - M)}{\kappa\sigma} > 1. \tag{57}$$

The next proposition records the tension between those two desirable properties, and a resolution.

Proposition 5.4 (Long run links between inflation, nominal rates and stability) *We have a positive long run link between inflation and nominal rates (56) and economic stability under passive monetary policy (57) if and only if $\alpha\zeta$ is large enough and agents are boundedly rational ($M < 1$), and prices are sticky enough. If $\alpha\zeta = 0$ (and "central bank guidance" has no impact) or $M = 1$, the two criteria cannot be simultaneously fulfilled.*

This proposition means that the system is determinate if enough agents follow the central bank's "inflation guidance", π_t^{CB} (i.e. if $\alpha\zeta$ is large enough). Intuitively, then, agents are "anchored" enough and the system has fewer multiple equilibria. Bounded rationality makes people's decision less responsive to the future (and in the old Keynesian model, to the past). As a result, it reduces the degree of complementarities, and we can more easily have only one equilibrium (this is quite a general point).

Leading old and new Keynesian models violate criterion (57), and allow for an unstable economy at the ZLB. The Old Keynesian model of Taylor (1999) has:[57] $\zeta = \beta^f = 0, \alpha = 1$. As a result, criterion (57) is violated if monetary policy is passive (as it is at the ZLB). However, criterion (57) tells us how to get stability in an Old Keynesian model: have $\zeta > 0$. If the economy is at the ZLB, *we avoid the deflationary spiral because of bounded rationality.* We need $M < 1$, and also $\alpha\zeta > 0$, i.e. current inflation is not very responsive to its past and future values.[58]

The traditional New Keynesian model has $M = 1$, so there is no stability (criterion (57) is violated). With $M < 1$, we can get stability. But to get "Fisher sign neutrality", ($\frac{di_\infty}{d\pi_\infty} > 0$), we need $\alpha\zeta > 0$.

[56] This is a controversial issue, as other authors have argued that the instability of the 1970s in the US was due to the Taylor criterion being validated. Within the model, the 1970s can be interpreted as a moment where agents do not believe the central bank enough, i.e. $\alpha\zeta$ is too low. This leads criterion (57) to be violated.

[57] It also replaces $i_t - \mathbb{E}_t \pi_{t+1}$ by $i_t - \pi_t$, and there is a x_{t-1} term in the IS equation, but that is a fairly immaterial difference.

[58] The Taylor model does feature a deflationary spiral, because it has $\zeta = 0$.

Hence, I conclude that the enrichment of this model is useful for both Old Keynesian and New Keynesian models.

Speculating somewhat more, this usefulness of "inflation guidance" may explain why central bankers these days do 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, which in turn can destabilize the economy. This reasoning relies on firms' bounded rationality.

5.3 Impulse Responses

A permanent shock to inflation. I assume that the central bank announces at time 0 an immediate, permanent, unexpected rise of 1% in the nominal rate and of its corresponding target inflation ($i_t = 1\%$ at all dates $t \geq 0$, and the central bank guidance is the corresponding long term target, $\pi_t^{CB} = \frac{1\%}{\frac{di_\infty}{d\pi_\infty}}$). Figure 5 shows the result. On impact, there is a recession: output and inflation are below trend. However, over time the 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 very hard to obtain in a conventional New Keynesian model. Cochrane (2015, p.1) summarizes the situation:[59]

> "If the Fed raises nominal interest rates, the [New Keynesian] model predicts that inflation will smoothly rise, both in the short run and long run. This paper presents a series of failed attempts to escape this prediction. Sticky prices, money, backward-looking Phillips curves, alternative equilibrium selection rules, and active Taylor rules do not convincingly overturn the result."

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 could well have negative welfare consequences.

A temporary shock to the interest rate. I now study a temporary increase of the nominal interest rate, $i_t = i_0 e^{-\phi t} > 0$ for $t \geq 0$. As the long run is not modified, I assume an inflation

[59] Cochrane (2015) needs to select a particular equilibrium, which leads to some controversy.

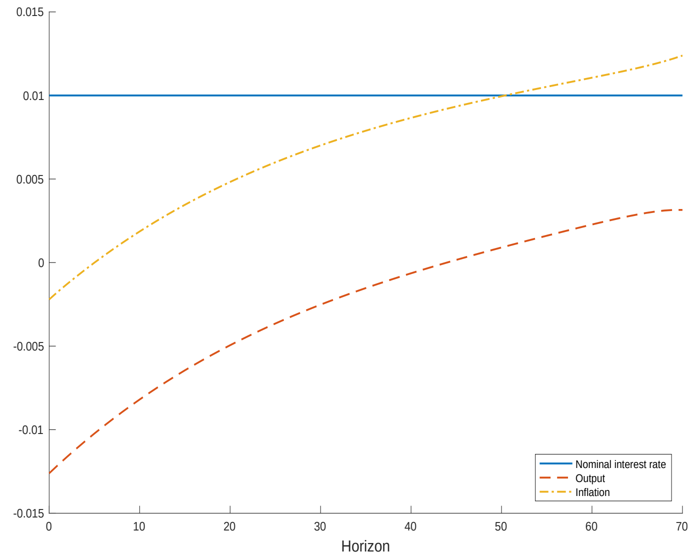

Figure 5: Impact of a permanent rise in the nominal interest rate. At time 0, the nominal interest rate is permanently increased by 1%. The Figure traces the impact on inflation and output.

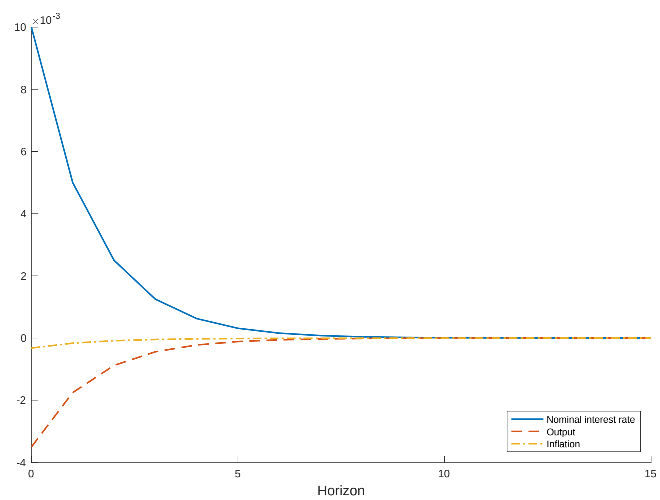

Figure 6: Impact of a temporary rise in the nominal interest rate. At time 0, the nominal interest rate is temporarily increased by 1%. The Figure traces the impact on inflation and output.

guidance of 0, $\pi_t^{CB} = 0$. Figure 6 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, i.e. setting $\alpha^f = 0$.

For most purposes, I recommend the basic model of Proposition 2.5. However, when the long run changes, the extension proposed in this section is useful. Substantively, it yields the insight that the equilibrium is determinate if agents are boundedly rational ($M < 1$) and they follow enough the "inflation guidance" by the government. Also, for empirical purposes, the extra backward term is helpful (Galí and Gerler 1999).

6 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 the model has good empirical support for its main non-standard elements. For instance, when Galí and Gertler (1999) estimate a Phillips curve, they estimate a partially myopic Phillips curve, which guided the numerical values in this paper. In the IS curve, the literature on the forward guidance puzzle, using a mix of market data and thought experiments, gives good evidence that we need $M < 1$, a main contention of the model. Finally, the notion that a higher interest rate lowers inflation in the short run (Keynesian effect), then raises it in the long run (classical Fisherian effect) is generally well-accepted, using again a mix of historical episodes and empirical evidence.

In conclusion, we have a theoretical model with empirical support for its non-standard features, that is also simple to use.

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 interesting conceptually: e.g. capital accumulation, a more frictional labor market, distortionary taxes, agents that are heterogeneous in wealth or rationality. The tractable approach laid out in this paper makes the exploration of those questions quite accessible.

Here I chose not to endogenize attention. This would not be too difficult (along the lines of Gabaix (2016)), but would complicate and enrich the policy analysis – e.g. attention to inflation depends on the Fed's aggressiveness in controlling inflation and vice-versa.

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 (12)), for firms (equation (21)), 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 (i.e. measuring the various m's in the economy).[60]

Surveys. It also suggests new questions for survey design. One would like to measure people's subjective model of the world – which, like in this model's agents, may not be an accurate model of the world. For instance, one could design surveys about people's understanding of the impulse-response 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 and Shleifer (2014)) rather than their whole causal model.[61] The parametrization 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 vs. a behavioral model.[62]

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.

---

[60] See progress on these issues in Coibon and Gorodnichenko (2015).

[61] 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.

[62] 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 "eight" parameter, and compare that to the rational and behavioral models.

7 Appendix: Behavioral 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 in 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:

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

The economy consists of a Dixit-Stiglitz continuum of firms. Firm i produces $q_{it} = N_{it}$ with unit productivity (there is no capital), and sets a price P_{it}. The final good is produced competitively in quantity $q_t = \left(\int_0^1 q_{it}^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}}$, and so that its price is:

$$P_t = \left(\int_0^1 P_{it}^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}. \tag{58}$$

A corrective wage subsidy $\tau = \frac{1}{\varepsilon}$ (financed by lump-sum transfers) ensures that there are no price distortions on average, so that the optimal price set by a firm (when it can reset its price) is $P_{it} = w_t$, so that price equals to marginal cost.[63]

The aggregate production function is $Y_t = N_t$. Calling GDP Y_t, the aggregate resource constraint is:

$$\text{Resource constraint: } Y_t = c_t + G_t = N_t. \tag{59}$$

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.

$$\text{Labor supply: } N_t^{\phi} = \omega_t c_t^{-\gamma}. \tag{60}$$

[63] This is well-known: $P_{it} = (1-\tau)\mu w_t = w_t$ with $\mu = \frac{\varepsilon}{\varepsilon-1}$.

The economy at time 1. Let us assume that the time-1 economy has flexible prices and no government consumption. Then, the real wage must equal productivity, $\omega_t = 1$. The labor supply equation (60) and $c_t = N_t$ give: $N_t^{\phi} = N_t^{-\gamma}$, so

$$N_1 = c_1 = 1.$$

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 maximizes $\max_{(c_t)_{t=0,1}} \sum_{t=0}^{1} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma}$ s.t. $\sum \frac{c_t}{R^t} = y_0 + \frac{y_1}{R_0}$. That gives

$$c_0 = b\left(y_0 + \frac{y_1}{R_0}\right) \tag{61}$$
$$b := \frac{1}{1+\beta}$$

with log utility.[64] Here b is the marginal propensity to consume (given the labor supply).[65]

Let us assume for now that government does not issue any debt nor consumes. Then aggregate income equal aggregate consumption: $y_t = c_t$. Hence,[66]

$$c_0 = b\left(c_0 + \frac{c_1}{R_0}\right) \tag{62}$$

which yields the Euler equation $\beta R_0 \frac{c_0}{c_1} = 1$. I use the consumption function formulation (62) 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 -1, not reactive to time-0 economic conditions, e.g. as in Mankiw and Reis

[64] 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 $\psi = 1$.

[65] 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 $c = N = 1$.

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

(2002)).[67] Other firms optimize freely their price, hence optimally choose a price

$$P_0^* = \omega_0 P_0 \tag{63}$$

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 (58), the aggregate price level is:

$$P_0 = \left(\theta \left(P_0^d\right)^{1-\varepsilon} + (1-\theta)\left(P_0^*\right)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} \tag{64}$$

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 ((59)–(60) and (62)-(64)). 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, $c_0 = N_0 = 1$.

In the model with sticky prices ($\theta > 0$), there is a one-dimensional continuum of real equilibria. It is the central bank who chooses the real equilibrium, by selecting the nominal interest rate, i.e., equivalently here, by choosing the real interest rate R_0.[68] 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 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 default, $y_1^d = c_1 = Y_1 = 1$.

But now suppose that (61) becomes:[69]

$$c_0 = b\left(y_0 + \frac{y_1^d + \bar{m}\hat{y}_1}{R_0}\right). \tag{65}$$

Suppose that the government consumes G_0 at 0, nothing at time 1, and makes a transfer $\mathcal{T}_t$ to

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

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

[69] Formally in terms of behavioral dynamic programming as in Gabaix (2016), this comes from the consumers maximizing:

$$(c_0, N_0) = \arg \underset{c_0, N_0 | m}{\text{smax}}\ u(c_0, N_0) + V\left(y_1^d + m_{\mathcal{T}}\widehat{y}_1 + R_0\left(\mathcal{T}_0 + \omega_0 N_0 - c_0\right)\right)$$

where V is the continuation value function. To make things very straightforward, consider that N_1 is fixed at 1. Then, $V(x) = u(x, 1)$.

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:

$$R_0 d_0 + \mathcal{T}_1 = 0. \tag{66}$$

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.[70] Hence, (65) 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 (66) we have:

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

and solving for c_0:

$$c_0 = \frac{b}{1-b}\left((1 - \bar{m})\, d_0 + \frac{Y_1}{R_0}\right). \tag{67}$$

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 (67). 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 (67) means that a positive transfer $d_0 = \mathcal{T}_0$ stimulates activity. If the government gives him $\mathcal{T}_0 > 0$ 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 stand on monetary policy to determine R_0. Here assume that the central bank does not change the interest rate R_0.[71]

[70] Note also that, 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 proceed of the time-0 government bonds to pay their taxes at time 1.

[71] 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).

Then, (67) implies that GDP ($Y_0 = c_0 + G_0$) changes as:

$$\frac{dY_0}{d\mathcal{T}_0} = \frac{b}{1-b}\left(1-\bar{m}\right). \tag{68}$$

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 7.1 *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:*

$$\frac{dY_0}{d\mathcal{T}_0} = \frac{b}{1-b}\left(1-\bar{m}\right)$$

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

$$\frac{dY_0}{dG_0} = 1 + \frac{b}{1-b}\left(1-\bar{m}\right).$$

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 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).[72] Equation (67) gives $\frac{dc_0}{dG_0} = \frac{b}{1-b}\left(1-\bar{m}\right)$, so that GDP, $Y_0 = c_0 + G_0$, has a multiplier

$$\frac{dY_0}{dG_0} = 1 + \frac{b}{1-b}\left(1-\bar{m}\right).$$

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 (60), 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 Ricardian equivalence fails), the multiplier $\frac{dY_0}{dG_0}$ is greater than 1. This is for the reason evoked in undergraduate textbooks: people feel richer, so spend more, which creates more demand. Here, we can assert that with good conscience – provided we allow for behavioral consumers.

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

Without Ricardian equivalence, the government consumption multiplier is greater than 1.[73] Again, this relies on monetary policy here 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 9.5.

The Phillips curve. Taking a log-linear approximation around $P_t = 1$, with $p_t = \ln P_t$, (64) 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.

$$\pi_0 = \frac{1-\theta}{\theta}\left(p_0^* - p_0\right) + \pi_0^d. \tag{69}$$

Via (63),

$$p_0^* - p_0 = \hat{\omega}_0 \tag{70}$$

where $\hat{\omega}_0 = \frac{\omega_0 - \omega_0^*}{\omega_0^*}$ is the percentage deviation of the real wage from the frictionless real wage, $\omega_0^* = 1$. Because of the labor supply condition (60), and $c_0 = N_0$, we have $\omega_0 = c_0^{\phi+\gamma}$. Therefore, $\hat{\omega}_0 = (\phi+\gamma)\hat{c}_0$. Hence (70) becomes $p_0^* - p_0 = (\phi+\gamma)\hat{c}_0$, and (69) yields:

$$\text{Phillips curve: } \pi_0 = \kappa \hat{c}_0 + \pi_0^d \tag{71}$$

[73] 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.

with $\kappa := \frac{1-\theta}{\theta}(\phi+\gamma)$. Hence, we obtain an elementary Phillips curve: increases in economic activity $\hat{c}_0$ lead to inflation. Inflation comes also from the automatic adjustment π_0^d.

To synthesize, we gather the results. Here $x_0 = \left(Y_0 - Y_0^d\right)/Y_0^d$ is the deviation of GDP from its frictionless value, $Y_0^d = 1$, while π_0 is the inflation between time -1 (the pre-time 0 price level) and time 0.[74] The deviations of $(c_0, G_0)$from trend are from the baseline of $(1, 0)$.

Proposition 7.2 (Two-period behavioral Keynesian model) *In this 2-period model, we have for time-0 consumption and inflation between periods 0 and 1:*

$$x_0 = \hat{G}_0 + b_d \hat{d}_0 - \sigma \hat{r}_0 \text{ (IS curve)} \tag{72}$$

$$\pi_0 = \kappa \hat{c}_0 + \pi_0^d \text{ (Phillips curve)} \tag{73}$$

where $\hat{G}_0$ is government consumption, $\hat{d}_0$ the budget deficit, $b_d = \frac{b}{1-b}(1-\bar{m})$ is the sensitivity to future deficits, $b = \frac{1}{1+\beta}$ is the marginal propensity to consume (given labor income) and $\hat{r}_0 = i_0 - \mathbb{E}\pi_1$ is the real interest rate between periods 0 and 1, and $\sigma = \frac{1}{R} = \beta$ with log utility.

This finished the derivation of the 2-period Keynesian model. The online appendix (Section 9.5) contains complements, including a discounted Euler equation.

8 Appendix: Additional Proofs

Proof of Proposition 2.3 The proof follows the steps and notations of Galí (2015, Section 3.3). I simplify matters by assuming constant return to scale ($\alpha = 0$ in Galí's notations). So, the nominal marginal cost at $t+k$ is simply ψ_{t+k}, not $\psi_{t+k|t}$.

Notations. When referring to equation 10 of Chapter 3 in Galí (2015), I write "equation (G10)" and do the same for (G11) and other equations. Lower-case letters denote logs. I replace the coefficient of relative risk aversion (σ in his notations) by γ (as in $u'(C) = C^{-\gamma}$).

Firms can reset their price with probability $1-\theta$. Generalizing Galí, firm i's optimal price p_t^* satisfies:

$$p_t^* - p_t = (1-\beta\theta) \sum_{k\geq 0} (\beta\theta)^k \mathbb{E}_t^{BR} \left[\psi_{t+k} - p_t\right]. \tag{74}$$

$$= (1-\beta\theta) \sum_{k\geq 0} (\beta\theta)^k \mathbb{E}_t^{BR} \left[p_{t+k} - p_t - \mu_{t+k}\right]. \tag{75}$$

[74]If the agent perceived only part of the change in the real rates, replacing R_0 by $(1-m_r) R_0^d + m_r R_0$ in (67), then the expression in (72) would be the same, replacing $\sigma = \frac{1}{R}$ by $\sigma = \frac{m_r}{R}$.

where I define $\mu_t := p_t - \psi_t$, the desired markup. This is a close cousin of the equation right before (G16).

Let us start by calculating:

$$\begin{aligned} H^0 : &= (1-\beta\theta)\sum_{k\geq 0}(\beta\theta)^k (p_{t+k} - p_t) = (1-\beta\theta)\sum_{k\geq 1}(\beta\theta)^k (\pi_{t+k} + ... + \pi_{t+1}) \\ &= (1-\beta\theta)\sum_{i\geq 1}\pi_{t+i}\sum_{k\geq i}(\beta\theta)^k = \sum_{i\geq 1}\pi_{t+i}(\beta\theta)^i = \sum_{i\geq 0}\pi_{t+i}(\beta\theta)^i 1_{i>0}. \end{aligned}$$

Hence

$$p_t^* - p_t = \sum_{k\geq 0}(\beta\theta)^k \mathbb{E}_t^{BR}\left[\pi_{t+k}1_{k>0} - (1-\beta\theta)\mu_{t+k}\right] \tag{76}$$

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

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

So:

$$\pi_t = \frac{1-\theta}{\theta}(p_t^* - p_t) \tag{77}$$

Plugging this in (75) gives:

$$\pi_t = \frac{1-\theta}{\theta}\mathbb{E}_t^{BR}\sum_{k\geq 0}(\beta\theta)^k\left(\pi_{t+k}1_{k>0} - \mu'_{t+k}\right) \tag{78}$$

where $\mu'_t := (1-\beta\theta)\mu_{t+k}$.

Next, I use the forward operator F ($Fy_t := y_{t+1}$), which allows to evaluate infinite sums compactly, as in (for some $\delta \in (0,1)$)

$$\sum_{k=0}^{\infty}\delta^k y_{t+k} = \sum_{k=0}^{\infty}\delta^k\left(F^k y_t\right) = \left(\sum_{k=0}^{\infty}\delta^k F^k\right)y_t = (1-\delta F)^{-1}y_t. \tag{79}$$

and define:

$$\rho := \beta\theta\bar{m} \tag{80}$$

Recall assumption (22)-(23) on $\mathbb{E}_t^{BR}[\pi_{t+k}]$ and $\mathbb{E}_t^{BR}[\mu_{t+k}]$. Plugging it in (78) gives

$$\begin{aligned}\pi_t &= \frac{1-\theta}{\theta}\mathbb{E}_t \sum_{k\geq 0} \rho^k \left(m_{f\pi}\pi_{t+k}1_{k>0} - m_{fx}\mu'_{t+k}\right) \\ &= \frac{1-\theta}{\theta}\frac{1}{1-\rho F}\left(m_{f\pi}\rho F \pi_t - m_{fx}\mu'_t\right)\end{aligned}$$

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

$$\pi_t = \beta M^f \mathbb{E}_t[\pi_{t+1}] - \lambda \mu_t \tag{81}$$

with

$$\begin{aligned}\beta^f &= \rho\left(1+\frac{1-\theta}{\theta}m_{f\pi}\right) = \beta\bar{m}\left(\theta + m_{f\pi}(1-\theta)\right) = \beta M^f && (82)\\ M^f &= \bar{m}\left(\theta + m_\pi^f (1-\theta)\right) && (83)\\ \lambda &= m_{fx}\frac{1-\theta}{\theta}(1-\beta\theta)\end{aligned}$$

The rest of the proof is as in Galí. The labor supply is still (60), $N_t^{\phi} = \omega_t C_t^{-\gamma}$, and as the resource constraint is $C_t = N_t$, $\omega_t = C_t^{-(\gamma+\phi)}$, i.e. $\hat{\omega}_t = -(\gamma+\phi)x_t$. This gives $\mu_t = \hat{\omega}_t$, i.e. (see Section 9.4 for a more general derivation):

$$\mu_t = -(\gamma+\phi)x_t. \tag{84}$$

Plugging this in (81), 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.

$$\begin{aligned}\kappa &= \bar{\kappa} m_x^f && (85)\\ \bar{\kappa} &= \left(\frac{1}{\theta}-1\right)(1-\beta\theta)(\gamma+\phi). && (86)\end{aligned}$$

Proof of Proposition 2.4 *Derivation of (26).* This comes naturally for the general formalism. Call $\boldsymbol{Z}_s = (B_s, d_s, d_{s+1}, d_{s+2}, ...)$ the state vector (more properly, the part of it that concerns deficits). Under the rational model, $\boldsymbol{Z}_{s+1} = \boldsymbol{H}\boldsymbol{Z}_s$ for a matrix $\boldsymbol{H}$: $(\boldsymbol{H}\boldsymbol{Z})(1) = \boldsymbol{Z}(1) + R\boldsymbol{Z}(2)$ and

$(\boldsymbol{HZ})(i) = \boldsymbol{Z}(i+1)$ for $i > 0$, where $\boldsymbol{Z}(i)$ is the $i-$th component of vector $\boldsymbol{Z}$. Under the cognitive discounting model simulated by the agent at time t, set $\boldsymbol{Z}_t^d = (B_t, 0, 0, ...)$ and the subjective model $\boldsymbol{Z}_s = \boldsymbol{Z}_t^d + \bar{m}\boldsymbol{H}\left(\boldsymbol{Z}_s - \boldsymbol{Z}_t^d\right)$. Hence the agent "sees" clearly the debt B_t, but more dimly the deficits d_t.We also have

$$T_s = -\frac{r}{R}B_s + d_s = \boldsymbol{e}^T\boldsymbol{Z}_s \text{ with } \boldsymbol{e}^T := \left(-\frac{r}{R}, 1, 0, 0, ...\right).$$

So,

$$\begin{aligned}
\mathbb{E}_t^{BR}[T_s] &= \mathbb{E}_t^{BR}\left[\boldsymbol{e}^T \cdot \boldsymbol{Z}_s\right] = \boldsymbol{e}^T \cdot \mathbb{E}_t^{BR}[\boldsymbol{Z}_s] = \boldsymbol{e}^T \cdot \left(\boldsymbol{Z}_t^d + (\bar{m}\boldsymbol{H})^{s-t}\left(\boldsymbol{Z}_s - \boldsymbol{Z}_t^d\right)\right) \\
&= \boldsymbol{e}^T \cdot \left(\boldsymbol{Z}_t^d + \bar{m}^{s-t}\mathbb{E}_t\left[\boldsymbol{Z}_s - \boldsymbol{Z}_t^d\right]\right) = -\frac{r}{R}B_t + \bar{m}^{s-t}\left(T_s + \frac{r}{R}B_t\right). \\
&= -\frac{r}{R}B_t + \bar{m}^{s-t}\left(d_s - r\sum_{u=t}^{\tau-1} d_u\right).
\end{aligned}$$

Derivation of (27). We have:

$$\begin{aligned}
x_t &= \frac{r}{R}k_t + \mathbb{E}_t^{BR}\left[\sum_{s\geq t}\frac{1}{R^{s-t}}b_y\left(\mathcal{T}_s + x_s\right)\right] = \frac{r}{R}B_t + b_y\sum_{s\geq t}\frac{\mathbb{E}_t^{BR}[\mathcal{T}_s]}{R^{s-t}} + b_y\sum_{s\geq t}\frac{\mathbb{E}_t^{BR}[x_s]}{R^{s-t}} \\
&= \frac{r}{R}B_t + \frac{r}{R}\sum_{s\geq t}\frac{-\frac{r}{R}B_t + m_y\bar{m}^{s-t}\left(d_s - r\sum_{\tau=t}^{s-1} d_\tau\right)}{R^{s-t}} + \mathbb{E}_t\left[\sum_{s\geq t}\frac{\bar{m}^{s-t}}{R^{s-t}}m_y b_y x_s\right] \\
&= \mathbb{E}_t\left[\sum_{s\geq t}\frac{\bar{m}^{s-t}}{R^{s-t}}m_y b_y\left(x_s + d_s - r\sum_{\tau=t}^{s-1} d_\tau\right)\right].
\end{aligned}$$

We see that the impact of B_t cancels out, a form of partial Ricardian equivalence. Old debt (B_t) does not make the agent feel richer. But a new deficit today (d_t) does.

This implies:

$$x_t = m_y\frac{r}{R}\left(x_t + Jd_t\right) + \frac{\bar{m}}{R}x_{t+1}$$

$$J := 1 - r\sum_{s\geq t+1}\frac{\bar{m}^{s-t}}{R^{s-t}} = 1 - r\frac{\frac{\bar{m}}{R}}{1-\frac{\bar{m}}{R}} = 1 - \frac{r\bar{m}}{R-\bar{m}} = \frac{R(1-\bar{m})}{R-\bar{m}}.$$

So, rearranging as in the derivation leading up to Proposition 2.2,

$$x_t = \frac{1}{R - m_y r}\left(Jrm_y d_t + \bar{m}x_{t+1}\right) = b_d d_t + \frac{\bar{m}}{R - m_y r}x_{t+1}$$

with $b_d = \frac{rm_y}{R-m_y r}\frac{R(1-\bar{m})}{R-\bar{m}}$. □

Proof of Proposition 3.1 We will use the following simple well-known fact.

Proposition 8.1 (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 8.1 indicates that the equilibrium is determinate iff: $|a| - 1 < b < 1$. Given 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 (38).

Proof of Proposition 3.2 Go back to (41), assuming the first best after the ZLB, so $\boldsymbol{z}_T = 0$. Then,

$$\boldsymbol{z}_0(T) = (\boldsymbol{A}_{ZLB} - \boldsymbol{I})^{-1}(\boldsymbol{A}_{ZLB}^T - \boldsymbol{I})\underline{\boldsymbol{b}}$$

In the traditional case, one of the eigenvalues of $\boldsymbol{A}_{ZLB}$ is greater than 1 in modulus. Then, $\|\boldsymbol{A}_{ZLB}^T\underline{\boldsymbol{b}}\| \to \infty$ (it's easy to verify that $\underline{\boldsymbol{b}}$ is not exactly the eigenvector corresponding to the root less than 1 in modulus). Hence, $\|\boldsymbol{z}_0(T)\| \to \infty$ as $T \to \infty$. Futhermore, this explosion is a recession: given that the entries of $\boldsymbol{A}_{ZLB}$ are positive, and those of $\underline{\boldsymbol{b}}$ are negative, each of the term in $\left(\boldsymbol{I} + \boldsymbol{A}_{ZLB} + ... + \boldsymbol{A}_{ZLB}^{T-1}\right)\underline{\boldsymbol{b}}$ is negative, hence, $\boldsymbol{z}_0(T)$ is has unboundedly negative inflation and output gap.

In the behavioral case however, all roots of $\boldsymbol{A}_{ZLB}$ are less than 1 in modulus. Hence, as $T \to \infty$, $\boldsymbol{z}_0(T) \to -(\boldsymbol{A}_{ZLB} - \boldsymbol{I})^{-1}\underline{\boldsymbol{b}}$, a finite value.

Proof of Proposition 4.3 The Lagrangian is

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

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 4.4 The central bank today takes as given its future actions, 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$-\pi_t - \Xi_t = 0$, which together yields $\pi_t = -\frac{\vartheta}{\kappa} x_t$. The explicit values of i_t is in Section 10.

Proof of Proposition 5.1 The IS curve is as in the baseline model. The only new term is the Phillips curve. I use (50) for $\mathbb{E}_t^{BR}[\pi_{t+k}]$ and (23) for $\mathbb{E}_t^{BR}[\mu_{t+k}]$. Plugging these in (78) gives, with $\rho = \beta\theta\bar{m}$

$$\begin{aligned} \pi_t &= \tfrac{1-\theta}{\theta} \left[\sum_{k\geq 1}\left((\beta\theta)^k - m_{f\pi}\rho^k\right)\right]\pi_t^d + \frac{1-\theta}{\theta}\mathbb{E}_t \sum_{k\geq 0}\rho^k\left(m_{f\pi}\pi_{t+k}1_{k>0} - m_{fx}\mu'_{t+k}\right) \\ &= \alpha^f \pi_t^d + \frac{1-\theta}{\theta}\frac{1}{1-\rho F}\left(m_{f\pi}\rho F \pi_t - m_{fx}\mu'_t\right) \\ \alpha^f &:= \frac{1-\theta}{\theta}\left(\frac{\beta\theta}{1-\beta\theta} - m_{f,\pi}\frac{\rho}{1-\rho}\right) \end{aligned} \tag{87}$$

Hence, multiplying by $1-\rho F$, $\pi_t = \beta^f \pi_{t+1} + \alpha^f (1-\rho F)\pi_t^d - \lambda\mu_t$ and using (84):[75]

$$\pi_t = \beta^f \mathbb{E}_t[\pi_{t+1}] + \alpha^f\left(\pi_t^d - \rho\pi_{t+1}^d\right) + \kappa x_t. \tag{88}$$

Proof of Proposition 5.3 The state vector is $\boldsymbol{z}_t = \left(x_t, \pi_t, \pi_t^d\right)$. Write the system as $\mathbb{E}_t \boldsymbol{z}_{t+1} = \boldsymbol{B}\boldsymbol{z}_t + a\pi_t^{CB}$, for a matrix $\boldsymbol{B}$. To study equilibrium multiplicity, we dispense with the forcing term $a\pi_t^{CB}$: indeed, the difference between two candidate equilibria will satisfy $\mathbb{E}_t \boldsymbol{z}_{t+1} = \boldsymbol{B}\boldsymbol{z}_t$. We can

[75]This implies, for the long run coefficient on inflation:

$$\begin{aligned} 1-\beta^f - (1-\rho)\alpha^f &= 1-\rho\left(1+\frac{1-\theta}{\theta}m_{f\pi}\right) - \frac{1-\theta}{\theta}\left(\frac{\beta\theta}{1-\beta\theta}(1-\rho) - m_{f,\pi}\rho\right) \\ &= (1-\rho)\left[1-(1-\theta)\frac{\beta}{1-\beta\theta}\right] = (1-\rho)\frac{1-\beta}{1-\beta\theta} = (1-\beta\theta\bar{m})\frac{1-\beta}{1-\beta\theta} \geq 1-\beta. \end{aligned}$$

write $\mathbb{E}_t \boldsymbol{z}_{t+1} = \boldsymbol{B} \boldsymbol{z}_t$, with (using $\chi := 1 - \rho$):

$$\boldsymbol{B} = \begin{pmatrix} \frac{\sigma\phi_x\beta^f + \beta^f + \kappa\sigma}{M\beta^f} & \frac{\sigma(\beta\phi_\pi - \alpha^f\eta\rho\chi - 1)}{M\beta^f} & \frac{\alpha^f((\eta-1)\rho+1)\sigma}{M\beta^f} \\ -\frac{\kappa}{\beta^f} & \frac{\alpha^f\eta\rho\chi+1}{\beta^f} & \frac{\alpha^f(-\eta\rho+\rho-1)}{\beta^f} \\ 0 & \eta\chi & 1-\eta \end{pmatrix} \tag{89}$$

Consider also the characteristic polynomial of $\boldsymbol{B}$, $\Phi(\Lambda) = \det(\Lambda\boldsymbol{I} - \boldsymbol{B})$ (with $\boldsymbol{I}$ the identity matrix), which factorizes as $\Phi(\Lambda) = \prod_{i=1}^{3}(\Lambda - \Lambda_i)$, where the Λ_i's are the eigenvalues of $\boldsymbol{B}$.

Inflation π_t^d is a predetermined variable, not a jump variable. Hence, for Blanchard-Kahn (1980) determinacy, $\boldsymbol{B}$ needs to have 1 eigenvalue less than 1 in modulus (corresponding to the predetermined variable π_t^d), and 2 greater than 1 (corresponding to the free variables x_t, π_t). This implies that $\Phi(1) > 0$, which is equivalent to (55), given that direct calculation of $\Phi(1) = \det(\boldsymbol{I} - \boldsymbol{B})$ gives:

$$\frac{\beta^f M}{\eta\kappa\sigma}\Phi(1) = \phi_\pi - 1 + \frac{1}{\kappa\sigma}\left(1 - (1-\zeta)\alpha - \beta^f\right)(1 - M + \sigma\phi_x).$$

References

Angeletos, George-Marios, and Chen Lian. 2016a. "Dampening General Equilibrium." Slides.

Angeletos, George-Marios, and Chen Lian. 2016b. "Forward Guidance without Common Knowledge." Working Paper, MIT.

Auclert, Adrien. 2015. "Monetary Policy and the Redistribution Channel." Working Paper, Princeton University.

Barro, Robert. 1974. "Are Government Bonds Net Wealth?" *Journal of Political Economy* 82: 1095-17.

Bilbiie, Florin O. 2008. "Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic." *Journal of Economic Theory* 140 (1): 162-96.

Blanchard, Olivier Jean, and Charles M. Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations." *Econometrica* 48 (5): 1305-11.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2016. "Diagnostic Expectations and Credit Cycles." Working Paper, Harvard University.

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

Brown, Jennifer, Tanjim Hossain, and John Morgan. 2010. "Shrouded Attributes and Information Suppression: Evidence from the Field." *Quarterly Journal of Economics* 125: 859-76.

Caballero, Ricardo. 1995. "Near-Rationality, Heterogeneity, and Aggregate Consumption." *Journal of Money, Credit and Banking* 27 (1): 29–48.

Caballero, Ricardo, and Emmanuel Farhi. 2015. "The Safety Trap." Working Paper, Princeton University.

Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2015. "Global Imbalances and Currency Wars at the ZLB." NBER Working Paper 21670.

Campbell, Jeffrey, Jonas Fisher, Alejandro Justiniano, and Leonardo Melosi. "Forward Guidance and Macroeconomic Outcomes Since the Financial Crisis." *NBER Macroeconomics Annual* 2016, vol. 31. University of Chicago Press, 2017.

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

Caplin, Andrew, Mark Dean, and Daniel Martin. 2011. "Search and Satisficing." *American Economic Review* 101 (7): 2899-22.

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

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

Cochrane, John. 2015. "Do Higher Interest Rates Raise or Lower Inflation?" Working Paper, Stanford University.

Cochrane, John. 2016a. "Comments on 'A Behavioral New-Keynesian Model' by Xavier Gabaix," Working Paper, Stanford.

Cochrane, John. 2016b. "Michelson-Morley, Occam and Fisher: The Radical Implications of Stable Inflation at Near-Zero Interest Rates," Working Paper, Stanford.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review* 105 (8): 2644-78.

Del Negro, Marco, Marc P. Giannoni, and Christina Patterson. 2015. "The Forward Guidance Puzzle." Working Paper, Federal Reserve Bank of New York.

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

Eggertsson, Gauti, and Neil Mehrotra. 2015. "A Model of Secular Stagnation". Working paper, Brown.

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

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

Evans, Martin. 1998. "Real Rates, Expected Inflation, and Inflation Risk Premia." *Journal of Finance* 53 (1): 187-218.

Eyster, Erik, Kristof Madarasz, and Pascal Michaillat. 2016. "Preferences for Fair Prices, Cursed Inference, and the Nonneutrality of Money." Working Paper, London School of Economics.

Farhi, Emmanuel, and Xavier Gabaix. 2015. "Optimal Taxation with Behavioral Agents." NBER Working Paper 21524.

Farhi, Emmanuel, and Ivan Werning. 2016. "The Power of Forward Guidance with Bounded Rationality and Incomplete Markets." Slides, Harvard University.

Fisher, Jonas DM. "On the Structural Interpretation of the Smets–Wouters 'Risk Premium" Shock'." *Journal of Money, Credit and Banking* 47 (2-3): 511-16.

Fuster, Andreas, Benjamin Hébert, and David Laibson. 2012. "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing." *NBER Macroeconomics Annual* 26: 1-48.

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

Gabaix, Xavier. 2016. "Behavioral Macroeconomics via Sparse Dynamic Programming." NBER Working Paper 21848.

Gabaix, Xavier, and David Laibson. 2002. "The 6D bias and the Equity Premium Puzzle." *NBER Macroeconomics Annual* 16: 257-312.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121: 505-40.

Gabaix, Xavier, and David Laibson. 2016. "Myopia and Discounting." Slides.

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

Galí, Jordi, and Mark Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44: 195-222.

Galí, Jordi, J. David López-Salido, and Javier Vallés. 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association* 5 (1): 227-70.

Garcia Schmidt, Mariana, and Michael Woodford. 2015. “Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis.” NBER Working Paper 21614.

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

Kandel, Shmuel, Aharon R. Ofer, and Oded Sarig. 1996. “Real Interest Rates and Inflation: An Ex-Ante Empirical Analysis.” *Journal of Finance* 51 (1): 205-25.

Kaplan, Greg, Ben Moll, and Giovanni L. Violante. 2016. “Monetary Policy According to HANK.” NBER Working Paper 21897.

Kiley, Michael. 2016. “Policy paradoxes in the New Keynesian model.” *Review of Economic Dynamics* 21: 1-15.

Kocherlakota, Narayana. 2016. “Fragility of Purely Real Macroeconomic Models.” NBER Working Paper 21866.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2015. “Business Cycle Dynamics under Rational Inattention.” *Review of Economic Studies* 82 (4): 1502-32.

Mankiw, N. Gregory. 2000. “The Savers-Spenders Theory of Fiscal Policy.” *American Economic Review Papers and Proceedings* 90 (2): 120-25.

Mankiw, N. Gregory, and Ricardo Reis. 2002. “Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve.” *Quarterly Journal of Economics* 117 (4): 1295-1328.

Mankiw, N. Gregory, and Matthew Weinzierl. 2011. “An Exploration of Optimal Stabilization Policy.” *Brookings Papers on Economic Activity* (Spring): 209-72.

McKay, Alisdair, Emi Nakamura, and Jon Steinsson. 2016. “The Power of Forward Guidance Revisited.” *American Economic Review* 106: 3133-58.

Nagel, Rosemarie. 1995. “Unraveling in Guessing Games: An Experimental Study.” *American Economic Review* 85 (5): 1313-26.

Reis, Ricardo. 2006. “Inattentive Consumers.” *Journal of Monetary Economics* 53: 1761-1800.

Sims, Christopher. 2003. “Implications of Rational Inattention.” *Journal of Monetary Economics* 50: 665–90.

Taubinsky, Dmitry, and Alex Rees-Jones. 2015. “Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment.” Working Paper, UC Berkeley.

Taylor, John B. 1999. “The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank.” *Journal of Monetary Economics* 43 (3): 655-79.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185: 1124–30.

Veldkamp, Laura. 2011. *Information Choice in Macroeconomics and Finance.* Princeton, NJ: Princeton University Press.

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

Werning, Iván. 2015. "Incomplete Markets and Aggregate Demand." Working Paper, MIT.

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

Woodford, Michael. 2010. "Robustly Optimal Monetary Policy with Near-Rational Expectations." *American Economic Review* 100: 274-303.

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

Woodford, Michael. 2012. "Inattentive Valuation and Reference-Dependent Choice." Working Paper, Columbia University.

Woodford, Michael. 2013. "Macroeconomic Analysis Without the Rational Expectations Hypothesis." *Annual Review of Economics* 5: 303–46.