

Working Paper

Behavioral Expectations in Nonlinear DSGE Models

Norges Bank Research

Author:
Junior Maih

Keywords:
Bounded rationality, cognitive
discounting, regime switching

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles på e-post: servicesenter@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på <http://www.norges-bank.no>

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail: servicesenter@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-8143 (online)

ISBN 978-82-8379-414-4 (online)

Behavioral Expectations in Nonlinear DSGE Models

Junior Maih*

Norges Bank

October 14, 2026[†]

Abstract

Cognitive discounting has become a workhorse device for bounded rationality in monetary economics. It is specified on perceived states or on selected expected terms and used mostly in *linearized* models, while the questions it is asked to answer — the effective lower bound, large shocks, risk, regime change — are intrinsically nonlinear. This paper introduces a variable-specific *behavioral-expectations operator*, defined directly on the nonlinear equilibrium conditions of DSGE models, including nonstationary and regime-switching ones. Before the rational expectation is taken, the operator contracts selected forward-looking variables toward reference points, in stated coordinates. The reference points may move over time or depend on the regime. At first order the operator reproduces Gabaix’s (2020) one-step discounting of expectations, except that each selected variable is discounted in every equation in which it appears. At second order and above it *discounts risk*: each variable’s k -th conditional cumulant of one-step deviations from its reference point, in the chosen coordinates (logs or levels), is scaled by the k -th power of its attention. This can turn divergent conditional expectations into convergent ones, a candidate *structural* complement to pruning. We illustrate the methodology by estimating a nonstationary nonlinear New Keynesian model with an occasionally binding lower bound on pre-2020 US data. With attention estimated variable by variable, the data decisively prefer cognitive discounting to rational expectations, by about 14 log points of marginal likelihood. Posterior mean attention is uneven across variables. Which restrictions the data accept depends on the model and the observables.

Keywords: bounded rationality; cognitive discounting; nonlinear DSGE; perturbation; regime switching; effective lower bound.

JEL: E31, E52, E70, C63.

*This draft should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. I thank Paolo Gelain, Johanna Krenz, Christian Matthes, and Greta Meggiorini for discussions during earlier stages of this project, and Xavier Gabaix for helpful exchanges. I also thank seminar and conference participants at the European Central Bank, the Bank of Lithuania, and the 32nd Annual Symposium of the Society for Nonlinear Dynamics and Econometrics (San Antonio) for useful comments, and an anonymous referee for constructive comments that improved the paper.

[†]First version: July 28, 2026.

Contents

1	Introduction	3
2	From rational to behavioral expectations	9
2.1	The idea in one variable	9
2.2	The operator	10
2.3	First order: Gabaix as a benchmark	12
2.4	Beyond first order: risk	14
2.5	Cumulants and finite beliefs	16
2.6	Trends and regimes	19
2.7	What the operator is	21
2.8	Using the operator	24
3	The quantitative laboratory	25
4	Results: beyond linear cognitive discounting	26
4.1	Cognitive discounting relaxes the Taylor principle	27
4.2	Higher-order dynamics: what attention does and does not stabilize	28
4.3	Impulse responses and business-cycle volatility	33
4.4	Forward guidance: the puzzle, and how attention resolves it	34
4.5	The cost of the lower bound	38
4.6	Fiscal policy: cognitive discounting raises the government-spending multiplier	40
4.7	Optimal policy in the nonlinear model, order by order	42
5	Estimation: methodology and results	47
5.1	Data, sample, and observables	47
5.2	Priors	47
5.3	Estimation strategy	48
5.4	Attention estimated variable by variable	48
5.5	How the data rank the expectation schemes	49
5.6	What identifies attention: the model and the observables	53
5.7	Robustness: a common-attention restriction	54
5.8	How the estimates compare with existing ones	59
5.9	A methodological note	60
6	Conclusion	60
A	The laboratory model	68
B	Where to find the details	75

1 Introduction

Cognitive discounting has become a workhorse device for bounded rationality in monetary economics. In [Gabaix \(2020\)](#), agents shrink their view of the future toward its steady state: expectations k periods ahead are attenuated by m^k for an attention parameter $m \in [0, 1]$. The device is disarmingly tractable and does real work — it resolves the forward-guidance puzzle, damps the power of far-ahead announcements, and improves the empirical fit of New Keynesian models. It sits alongside a broader literature on inattention and finite planning horizons ([Maćkowiak and Wiederholt, 2009](#); [Angeletos and Lian, 2018](#); [Woodford, 2019](#)) that has been developed mostly in a *linearized* world (the nonlinear exceptions are discussed below). So has the applied program built on it: cognitive discounting has been estimated in linearized models ([Andrade et al., 2019](#); [Ilabaca et al., 2020](#); [Meggiorini, 2023](#); [Afsar et al., 2024](#); [Hohberger et al., 2024](#)) and extended — at first order — to heterogeneous agents ([Pfäuti and Seyrich, 2022](#)), open economies ([Kolasa et al., 2025](#)), optimal policy ([Benchimol and Bounader, 2023](#)), and fiscal policy ([Lubik and Marzo, 2021](#)).

This creates a mismatch between the tool and the questions it is used to answer. It is specified on perceived states or on selected expected terms and used mostly in *linearized* models, while the questions it is asked to answer — the effective lower bound, large shocks, risk, regime change — are intrinsically nonlinear. They are precisely the phenomena at which a first-order approximation is least reliable, and the models that policy institutions use to study them are nonlinear, nonstationary, and regime-switching. The behavioral tool and much of the economics it is asked to produce therefore do not live in the same place.

Two qualifications sharpen rather than weaken this point. First, not every application needs nonlinearity: the forward-guidance puzzle, for instance, arises already in a linear model away from the bound, and there the linear device is adequate. Second, it is far from obvious that *inattention* is strongest in the episodes where nonlinearity bites. Attention is plausibly scarce in tranquil times and sharpens when volatility rises, when inflation is high, or when the regime changes. Information rigidities in survey forecasts fell when macroeconomic volatility was high and rose during the Great Moderation ([Coibion and Gorodnichenko, 2015](#)), and attention to inflation rises with its level ([Pfäuti, 2025](#); [Bracha and Tang, 2025](#); [Weber et al., 2025](#)). Cognitive discounting may then be, first and foremost, a description of normal times. But this is itself a statement about how attention varies with the state of the economy. It is naturally examined in a framework that is defined in both states — in and out of the effective lower bound (ELB), in calm and in volatile periods — and that allows attention to differ between them. The operator proposed here is such a framework ([Proposition 4](#) and [Remark 7](#)); whether attention rises or falls in bad times is then a question for the model and

the data rather than an assumption.

This paper closes that gap. It confronts two specific problems with the way cognitive discounting is currently formulated.

Gap 1: discounting is specified on states or on expected terms. In Gabaix (2020) agents perceive next period’s state through a contraction of its law of motion toward the steady state. The perceived transition is specified nonlinearly and then linearized, which delivers the attenuation m^k ; like any perceived transition, it also induces an expectation operator by integration. Sparse dynamic programming extends the approach to nonlinear decisions (Gabaix, 2016a). Hirose et al. (2024) instead discount the expectations of selected forward-looking terms toward their steady-state values, in a nonlinear model with growth and the zero lower bound. Neither route contracts selected *variables*, in stated coordinates and toward stated reference points, identically wherever those variables appear in the nonlinear equilibrium conditions. This paper characterizes what that construction implies for *risk*. Because a linearized model exhibits certainty equivalence, the linearized device is silent about a channel that ought to be central to behavioral macro: if inattentive agents shrink their view of the future, do they also under-react to future *uncertainty*? The question cannot be posed in a linearized formulation.

Gap 2: “what to discount” is ad hoc. Even within linear models the textbook device is ambiguous once a variable is forward-looking in *several* equations, or when *several* forward-looking terms share one equation — for example a wage Phillips curve in $\mathbb{E}_t \lambda_{t+1}$, $\mathbb{E}_t p_{t+1}$, $\mathbb{E}_t y_{t+1}$, $\mathbb{E}_t \pi_{t+1}^w$. Which of these does the agent discount, and by how much? Applied loosely, the device becomes a free parameter attached to equations rather than to beliefs. The ambiguity is not algebraic: with one coefficient and coherent reference values, discounting an expected term is linear in that term, so splitting it into additive pieces changes nothing. It lies in which economic objects are discounted, with which coefficients and reference values.

This matters most when the model is estimated. A discount factor attached to one equation, or to one forward term but not another, is an additional parameter that the likelihood can use for purposes unrelated to expectations: to add persistence the model otherwise lacks, to offset a misspecified shock process, or to trade off against the volatility of a shock that moves the same expectations. The estimate then improves fit, but it no longer measures what agents believe, and estimates from different models cease to be comparable. Two features of the approach below address this. The discount is attached to *variables*, and a discounted variable is treated identically in every equation in which it appears, so that ρ

has a single interpretation as a shrinkage of the perceived law of motion of that variable. And the estimation reports which attentions the data identify, and checks whether attention is confounded with the volatility of the shocks that move the same expectations (Remark 8, §5.6 and §5.7.3).

The consequence is that, in a serious estimated nonlinear model, one cannot ask the questions that motivate the behavioral program in the first place: *which* variables agents discount, by *how much*, and whether *attention itself moves with the state of the economy* — for instance, whether agents pay less attention to the future precisely when policy is constrained at the lower bound.

The empty cell. These literatures exist separately but rarely together. Work on bounded rationality and inattention in macroeconomics is predominantly linear or linearized (Gabaix, 2020; Angeletos and Lian, 2018; Maćkowiak and Wiederholt, 2009; Woodford, 2019): Gabaix’s perceived transition is nonlinear before it is linearized, and sparse dynamic programming (Gabaix, 2016a) is a notable nonlinear exception. The forward-guidance literature treats discounting as a linear fix (McKay et al., 2016; Del Negro et al., 2023). The nonlinear and ELB literatures are, with two instructive exceptions, rational-expectations (Fernández-Villaverde et al., 2015; Nakata, 2017; Aruoba et al., 2018), and so is the regime-switching DSGE literature (Hamilton, 1989; Davig and Leeper, 2007; Bianchi, 2013; Foerster et al., 2016; Binning and Maih, 2017). The exceptions mark the frontier. Hirose et al. (2024) estimate a small nonlinear New Keynesian model with stochastic growth and the ZLB under Gabaix-style discounting, solved globally, with separate discounts for households and firms. The discount is applied to the expectation of each selected forward-looking *term* of that model’s recursions: the expected term is shrunk toward its steady-state value by a common factor, so what must be chosen is which terms are discounted (the ambiguity of Gap 2). The operator proposed here discounts *variables* instead; with matching common attention and reference values the two coincide at first order, and their nonlinear risk loadings generally differ (§5.7.4). Ascari et al. (2023) show that behavioral discounting can restore the *existence* of equilibrium at the lower bound in piecewise-linear models. This paper contributes a variable-specific transformation defined directly on the nonlinear equilibrium conditions, with explicit choices of variables, coordinates and reference points, and the same treatment of a selected variable in every equation. It nests the linearized device at first order, characterizes what the contraction does to perceived risk, and extends to growing and regime-switching economies through moving or regime-specific reference points and regime-dependent attention.

The approach. This paper introduces a variable-specific behavioral-expectations operator, defined directly on the nonlinear equilibrium conditions of DSGE models, including non-stationary and regime-switching ones. Before the rational expectation is taken, the operator contracts selected forward-looking variables toward reference points, in stated coordinates. The reference points may move over time or depend on the regime. The single operator acts on the nonlinear equilibrium conditions $\mathbb{E}_t f_\theta(X_{t+1}, X_t, X_{t-1}, \varepsilon_t) = 0$ by

$$\mathbb{E}_t^{\text{BR}} f_\theta(X_{t+1}, \cdot) = \mathbb{E}_t f_{\{\theta, \rho_X\}}(\bar{X} \odot (X_{t+1} \odot \bar{X})^{\circ \rho_X}, \cdot), \quad \hat{x}_{t+1} \mapsto \rho_X \odot \hat{x}_{t+1},$$

where $\hat{x} \equiv \log X - \log \bar{X}$: the contraction is in logs, toward the reference point $\bar{X}$, by a variable-specific attention parameter $\rho_X \in [0, 1]$. The contraction is applied to the full nonlinear system — it requires neither linearity nor stationarity — and it disciplines “what to discount”: a variable selected for behavioral treatment is treated identically wherever it appears. Rational expectations ($\rho_X = \mathbf{1}$) and full myopia ($\rho_X = \mathbf{0}$) are the two corners.

Defining the operator on the model as written makes computable several questions that the linear device cannot pose. How do inattentive agents perceive future risk? How does attention shape higher-order dynamics and the boundedness of their solutions? How does it change optimal policy under commitment and under discretion? What does it imply at an occasionally binding lower bound? And how much attention do the data support when the nonstationary model is estimated directly? These are the applications of Sections 4 and 5.

Contributions. The paper makes four contributions.

1. *A variable-specific operator* defined directly on the nonlinear equilibrium conditions, with explicit choices of variables, coordinates and reference points, applied identically wherever a selected variable appears, well defined at any perturbation order and for growing (balanced-growth) and regime-switching models (§2). It differs from contracting the perceived state (Gabaix, 2020) and from discounting selected expected terms (Hirose et al., 2024), and with matching common attention and reference values it agrees with the latter at first order.
2. *Theory* (§2). At first order the operator reproduces Gabaix’s (2020) one-step discounting of expectations, except that each selected variable is discounted in every equation in which it appears. At second order and above it *discounts risk*: each variable’s k -th conditional cumulant of one-step deviations from its reference point, in the chosen coordinates (logs or levels), is scaled by the k -th power of its attention. This can turn divergent conditional expectations into convergent ones, a candidate *structural* complement to pruning. In detail, at first order the operator nests Gabaix’s attenuation

of expectations when the common discount covers the state (Proposition 1). It also reproduces his discounted curves, except where he leaves the expectation of a selected variable undiscounted; in the laboratory model, that is expected inflation in the real interest rate (Remark 2). At order ≥ 2 , the loading on each conditional variance and covariance of future variables' deviations from a fixed reference point is scaled by the product of their attentions. With common attention, the whole risk term is scaled by ρ^2 . Under certainty equivalence this term, and with it the under-reaction to uncertainty, is identically zero (Proposition 2, new). More generally, the operator scales each variable's k -th one-step-ahead conditional cumulant of the chart deviation by the k -th power of its attention, and joint cumulants by the product of the attentions involved, so behavioral expectations can be finite where rational ones diverge (Proposition 3 and Corollary 2, new). This is a possible mechanism, checked numerically in the laboratory model, by which the operator can stabilize divergent higher-order dynamics as a structural complement to pruning (Remark 5). Finally, it extends to regime switching with regime-dependent reference points and state-dependent attention (Proposition 4).

3. *Quantitative lessons* (§4) in one estimated nonlinear, nonstationary New Keynesian model with an occasionally binding lower bound. Cognitive discounting relaxes the Taylor principle (shown in the three-equation model), weakens forward guidance away from the bound (at given structural parameters), and raises the government-spending multiplier — lessons that carry over from *linear* cognitive discounting (Gabaix, 2020). Beyond first order, in repeated simulations, low enough attention removes the explosions of un-pruned higher-order solutions that are driven by persistence, except at the highest persistence we consider; explosions driven by large shocks are not removed in general. The operator also lets us pose welfare-optimal (Ramsey) and discretionary policy in a stationary variant of the model, where cognitive discounting slightly narrows the gain from commitment. With a common attention, a permanent peg at the occasionally binding lower bound is locally determinate only below a threshold $\rho^* \in [0.720, 0.736]$ — a first-order property, in the spirit of Angeletos and Lian (2023), that the operator carries to the kinked model.
4. *Empirics* (§5): because the operator is defined on the estimable model, the behavioral question can be put to the data. We illustrate the methodology by estimating a non-stationary nonlinear New Keynesian model with an occasionally binding lower bound on pre-2020 US data. The estimation keeps the lower bound in the likelihood. With attention estimated variable by variable, the data decisively prefer cognitive discounting to rational expectations, by about 14 log points of marginal likelihood. Posterior mean

attention is uneven across variables. Which restrictions the data accept depends on the model and the observables. On these 1985–2019 data, cognitive discounting beats rational expectations under every attention structure we estimate, by 12.8 to 14.6 log points (§5.5). The two reference-inflation attentions have lower posterior means (0.45 and 0.52) than the others (between 0.76 and 0.91) and wide posteriors. The data give positive evidence for freeing them jointly rather than imposing rational expectations on both, a distinction that the common and grouped restrictions considered here cannot reveal. The data positively favor attention by variable over block-specific and common restrictions. A two-group households/firms restriction fits about as well in this model, a result learned by comparing it with the full variable-by-variable model. The estimation illustrates the method in one model with one set of observables; the operator is designed for richer models, where such restrictions become testable rather than assumed.

The operator is implemented in RISE (Maih, 2015), so the object analyzed in the theory and the object solved and estimated in the applications are the same.

Preview of the results. Three findings stand out. First, the discounting of risk is exact and easy to see. In a consumption Euler equation, uncertainty about future consumption lowers the required return. This precautionary term shrinks with the *square* of attention, holding the conditional consumption distribution fixed: at $\rho = 0.8$ it is 64 percent of its rational-expectations magnitude (Online Appendix OA-B.5). Second, in the laboratory model’s simulations, low enough attention removes the explosions of higher-order solutions whose divergence comes from persistence. Once shock persistence reaches 0.99, un-pruned rational simulations explode at third order, while under cognitive discounting our criterion detects no explosion for $\rho \leq 0.95$. Lower attention is needed as persistence rises further ($\rho \leq 0.70$ at 0.998), and at 0.999 no value on our grid suffices (§4.2). Cognitive discounting does not in general remove explosions driven by large shocks, and at second order some intermediate values of attention produce explosions where the rational solution does not, so it complements pruning rather than replacing it. Third, with attention estimated variable by variable the data decisively prefer cognitive discounting, which beats rational expectations by about 14 log points of marginal likelihood. Posterior mean attention is uneven across variables: between 0.76 and 0.91 per quarter for consumption, price inflation, marginal utility and wage inflation, around the value Gabaix (2020) calibrates, and near one half, with wide posteriors, for the two reference inflation rates. For these two, the data positively favor freeing the attentions over imposing rational expectations on both (§§5.4–5.5). In the applications of Section 4, at given structural parameters, forward guidance is weaker

in normal times. At the lower bound, however, the effect of a calendar peg keeps growing with its length over the horizons computed at the restricted estimate, a common attention of 0.897. With a common attention, a permanent peg is locally determinate only below $\rho^* \in [0.720, 0.736]$ (§4.4); with attention by variable, determinacy depends on the whole attention vector. Alongside these, the familiar lessons of linear cognitive discounting — a relaxed Taylor principle, weaker forward guidance at given structural parameters, a larger fiscal multiplier — carry over to the nonlinear model, with magnitudes disciplined by the model’s structure.

Roadmap. Section 2 introduces the operator with a one-variable example, defines it, states its properties at first order, for risk, and for the whole one-step distribution, extends it to trends and regimes, relates it to other models of expectation formation, and closes with the steps for using it. Section 3 describes the nonlinear New Keynesian laboratory model. Section 4 reports the quantitative results. Section 5 estimates attention variable by variable and compares it with rational expectations and with restrictions of the operator. Section 6 concludes. Appendix A lists the laboratory model’s equilibrium conditions and priors, and Appendix B says where each proof and technical detail is. The Online Appendix collects them, grouped by question: beliefs and foundations, risk and integrability, trends and regimes, the three-equation algebra, estimation details, and lower-bound results.

2 From rational to behavioral expectations

This section teaches the operator before it audits it. Subsection 2.1 shows the idea on a single variable, and §2.2 defines the operator for a whole model. Sections 2.3–2.5 state what it does at first order, to risk, and to the whole one-step distribution. Section 2.6 covers trends and regimes. Section 2.7 says what the operator is, and is not, as a model of beliefs, and §2.8 collects the steps for using it. Each result is stated with the conditions that make it true; proofs, full regularity conditions and the formal perceived-law results are in the Online Appendix, and Appendix B lists where each one is.

2.1 The idea in one variable

What does it mean to discount a *variable* rather than an *expectation*, and what changes beyond first order? Take a household forecasting next period’s marginal utility of consumption. Let $c \equiv \log C_{t+1} - \log \bar{C}$ be the log deviation of next period’s consumption from a reference point normalized to $\bar{C} = 1$, so that with log utility the object forecast is $g(c) = e^{-c} = 1/C_{t+1}$.

Write $c = \varsigma \tilde{c}$, with ς the scale of fluctuations, and let $\mu \equiv \mathbb{E}_t c$ and $v \equiv \text{Var}_t(c)$. Assume $\mathbb{E}_t[|\tilde{c}|^3 e^{\bar{c}|\tilde{c}|}] < \infty$ for some $\bar{c} > 0$ (the condition of Proposition 2); finite mean and variance alone would not make $\mathbb{E}_t e^{-c}$ finite or its Taylor remainder of third order. An inattentive agent can be modeled in two ways:

- *discount the variable*: contract c toward its reference point by a factor $\rho \in [0, 1]$, then forecast, $\mathbb{E}_t e^{-\rho c}$;
- *discount the term*: shrink the expected deviation of the term from its reference value $g(0) = 1$ by a factor $M \in [0, 1]$, $1 + M(\mathbb{E}_t e^{-c} - 1)$.

To second order in ς , using $\mathbb{E}_t c^2 = \mu^2 + v$,

$$\text{variable rule: } \mathbb{E}_t e^{-\rho c} = 1 - \rho \mu + \frac{\rho^2}{2} (\mu^2 + v) + O(\varsigma^3), \quad (1)$$

$$\text{term rule: } 1 + M(\mathbb{E}_t e^{-c} - 1) = 1 - M\mu + \frac{M}{2} (\mu^2 + v) + O(\varsigma^3). \quad (2)$$

With $M = \rho$ the two rules agree at first order: both shrink the expected deviation by the same factor, which is the familiar discounting of expectations. They differ at second order. The variable rule loads the variance with weight ρ^2 , the term rule with weight M ; at $\rho = M = 0.8$ an inattentive agent responds to the variance of future consumption with weight 0.64 under the first and 0.8 under the second. If c is conditionally Gaussian the variable rule is exact in logs, $\log \mathbb{E}_t e^{-\rho c} = -\rho \mu + \frac{1}{2} \rho^2 v$: the mean is scaled by ρ and the variance by ρ^2 . The term rule acts only on the expected deviation of the term.

Two conditions are built into this example. The distribution of c and the reference point are held fixed; in equilibrium both change with attention, so the example ranks neither welfare nor fit (the two rules are compared in the estimated model in §5.7.4). And there is a single variable, so attention is common; with several variables and unequal attentions the loadings are scaled covariance by covariance (§2.4). The rest of the section generalizes the example: to a whole system of equations (§2.2), to Gabaix's first-order benchmark (§2.3), to every conditional variance and covariance (§2.4), and to every cumulant of the one-step distribution (§2.5).

2.2 The operator

Consider a DSGE model whose equilibrium conditions are

$$\mathbb{E}_t f_\theta(X_{t+1}, X_t, X_{t-1}, \varepsilon_t) = 0, \quad (3)$$

where X_t is the n -vector of endogenous variables, ε_t the exogenous shocks, θ the parameters and $\mathbb{E}_t$ the rational expectations operator.

Definition 1 (Behavioral expectations). *For each variable X_i fix a chart χ_i : $\chi_i = \log$ if X_i is strictly positive (levels of quantities and prices, including balanced-growth variables), and $\chi_i = \text{id}$ if X_i is real-valued (rates and gaps, e.g. net inflation, which may be negative). Let $\bar{X}$ denote the (possibly regime-specific, possibly balanced-growth) reference point of X , and let $\rho_X \in [0, 1]^n$ be a vector of variable-specific cognitive discounting parameters. The behavioral expectations operator $\mathbb{E}_t^{\text{BR}}$ replaces every one-step-ahead variable by the contraction of its chart deviation toward the reference point before the (rational) expectation is taken:*

$$\chi_i(X_{i,t+1}) \mapsto \chi_i(\bar{X}_i) + \rho_{X_i} [\chi_i(X_{i,t+1}) - \chi_i(\bar{X}_i)]. \quad (4)$$

In the log chart this is the geometric shrinkage

$$\mathbb{E}_t^{\text{BR}} f_\theta(X_{t+1}, X_t, X_{t-1}, \varepsilon_t) \equiv \mathbb{E}_t f_{\{\theta, \rho_X\}}(\bar{X} \odot (X_{t+1} \oslash \bar{X})^{\circ \rho_X}, X_t, X_{t-1}, \varepsilon_t), \quad (5)$$

equivalently, in logs,

$$\log(\bar{X} \odot (X_{t+1} \oslash \bar{X})^{\circ \rho_X}) = \log \bar{X} + \rho_X \odot [\log X_{t+1} - \log \bar{X}]; \quad (6)$$

in the level chart it is the affine shrinkage $X_{t+1} \mapsto \bar{X} + \rho_X \odot (X_{t+1} - \bar{X})$. $\rho_X = \mathbf{1}$ recovers rational expectations; $\rho_X = \mathbf{0}$ is full myopia ($X_{t+1} \mapsto \bar{X}$).

For example, next period's consumption, a positive level, is perceived as $\bar{C}^{1-\rho} C_{t+1}^\rho$, a geometric average of the realization and the reference point; next period's net inflation, a rate, is perceived as $\bar{\pi} + \rho(\pi_{t+1} - \bar{\pi})$. Throughout, $\hat{x}_t \equiv \chi(X_t) - \chi(\bar{X})$ denotes the vector of chart deviations from the reference point — log deviations for log-chart variables, level deviations otherwise — and $D_\rho \equiv \text{diag}(\rho_X)$. In this notation the shrinkage (4) reads simply $\hat{x}_{t+1} \mapsto D_\rho \hat{x}_{t+1}$. Because this map is linear and $\mathbb{E}_t$ is linear, the operator attenuates the one-step conditional mean of the chart deviation exactly,

$$\mathbb{E}_t^{\text{BR}} \hat{x}_{t+1} = D_\rho \mathbb{E}_t \hat{x}_{t+1}, \quad (7)$$

and therefore, to first order around the reference point, the forward block of every equilibrium condition. Everything beyond first order in this section is about what (7) leaves out.

Joint feasibility. The contraction keeps each variable in its own domain (positive levels stay positive), but it need not keep the *vector* in the economically meaningful domain of

f . With habit, for instance, the perceived surplus can cross the singularity of log utility although every level stays positive: with reference consumption 1, $C_t = C_{t+1} = 2$, habit $b = 0.8$ and $\rho = \frac{1}{2}$, the true future surplus is $C_{t+1} - bC_t = 0.4 > 0$ but the perceived one is $\sqrt{2} - 1.6 < 0$. Readings of the operator as a household's perceived problem are therefore local: they require the contracted states to remain in the admissible domain of f , which holds in a neighborhood of a feasible steady state, the setting of the perturbation solutions used in this paper. Integrability of f alone does not restore it.

This chart assignment is as implemented in RISE, so the operator analyzed here and the operator the computations solve are the *same object*, chart by chart — a correspondence we verified numerically.

Remark 1 (Variable-specific, system-wide discounting). *The operator acts on the system and on selected variables. First, discounting is attached to variables, whereas Gabaix's operator contracts the perceived state and term-wise rules discount expected terms, in both cases with agent-specific coefficients. Second, a variable selected for behavioral treatment is treated identically everywhere it appears. This disciplines the two configurations the basic framework leaves open: several forward-looking variables in one equation, and the same forward-looking variable in several equations. Third, nothing in (5) requires linearity or stationarity.*

2.3 First order: Gabaix as a benchmark

What does the operator do in a linearized model, and how does it relate to the cognitive discounting of Gabaix (2020)? Iterating (7) answers both.

Proposition 1 (Gabaix nesting at first order). *Suppose the behavioral equilibrium exists, is determinate, and has the first-order law of motion $\hat{x}_{t+1} = A\hat{x}_t + B\varepsilon_{t+1}$, where A is the transition of the behavioral equilibrium itself and $\hat{x}$ stacks all state variables, exogenous processes included. Let agents apply $\mathbb{E}^{\text{BR}}$ at every date and form k -step forecasts by composing it, with this transition as the perceived law.*

(i) For every $k \geq 1$,

$$\mathbb{E}_t^{\text{BR}} \mathbb{E}_{t+1}^{\text{BR}} \cdots \mathbb{E}_{t+k-1}^{\text{BR}} \hat{x}_{t+k} = (D_\rho A)^k \hat{x}_t. \quad (8)$$

(ii) With a common discount on every variable of the state, exogenous processes included, $\rho_X = m\mathbf{1}$, $D_\rho = mI$ commutes with A and

$$\mathbb{E}_t^{\text{BR}} \mathbb{E}_{t+1}^{\text{BR}} \cdots \mathbb{E}_{t+k-1}^{\text{BR}} \hat{x}_{t+k} = m^k \mathbb{E}_t \hat{x}_{t+k}, \quad k \geq 1, \quad (9)$$

which is exactly Gabaix's cognitive discounting of expectations.

(iii) *Applied to the textbook three-equation New Keynesian model with the output gap x and inflation π selected at a common m , the operator delivers the discounted IS and Phillips curves*

$$x_t = m \mathbb{E}_t x_{t+1} - \sigma(i_t - m \mathbb{E}_t \pi_{t+1} - r_t^n), \quad (10)$$

$$\pi_t = \beta m \mathbb{E}_t \pi_{t+1} + \kappa x_t. \quad (11)$$

Proposition 1 says that, to first order, the operator is a generalization of Gabaix’s device: a matrix D_ρ in place of the scalar m . Two conditions carry its part (ii). First, the common discount must cover the *state*, exogenous processes included. In the three-equation application only x and π are discounted, so the exogenous state r^n propagates uncontracted and agents’ composed forecasts are $m \mathbb{E}_t x_{t+k}$: the attenuation does not compound (Online Appendix, Corollary OA-A.1). The horizon discount m^k that Gabaix uses to summarize forward guidance then arises in equilibrium instead, by iterating (10) forward, $x_t = -\sigma \sum_{k \geq 0} m^k \mathbb{E}_t [i_{t+k} - m \pi_{t+k+1} - r_{t+k}^n]$, for $m < 1$ and bounded conditional means (at $m = 1$ the terminal condition $\lim_{K \rightarrow \infty} \mathbb{E}_t x_{t+K} = 0$ must be imposed). Second, composed forecasts require the perceived transition to be specified at the off-equilibrium points that the contracted chain reaches; the equilibrium law alone does not pin it down there (Online Appendix §OA-A.3). The proposition is a first-order statement: beyond first order composed forecasts are in general not $(D_\rho A)^k$ (Proposition OA-A.2). The proof is in Online Appendix §OA-A.1.

Remark 2 (What is, and is not, nested). *The nesting is of Gabaix’s cognitive-discounting operator: his attenuation of expectations, m^k at horizon k , is reproduced exactly when every variable of the state, exogenous processes included, is discounted (Proposition 1). When only some variables are discounted, as in the three-equation application, the horizon discount arises in equilibrium through the forward solution rather than in agents’ composed forecasts. The micro-founded Phillips curve is nested as well. Gabaix (2020) reports a firm-level coefficient $M^f = m[\theta + (1 - \theta)\frac{1-\beta\theta}{1-m\beta\theta}]$, which differs from m (it is strictly below m for every $0 < m < 1$; both are zero at $m = 0$). Benchimol and Bounader (2023) apply the term structure of behavioral expectations to deviations from the steady state rather than to levels, consistently with Gabaix’s own lemma on that term structure. Absent additional firm-specific inattention to inflation, this yields a firm coefficient equal to m . With that coefficient, Gabaix’s Phillips curve is exactly (11). The one remaining difference is the IS curve: Gabaix leaves the inflation expectation in the real rate undiscounted, whereas the operator, applied to π , discounts it wherever it appears — the uniform treatment that Gap 2 calls for. (Gabaix notes that consumer inattention to inflation would discount this term as well.) The*

experiments of §4 therefore use Gabaix’s curves up to this single term.

Remark 3 (Beyond a scalar discount). *With variable-specific ρ_X , D_ρ and A need not commute, and the agents’ composed forecasts $(D_\rho A)^k \hat{x}_t$ are richer than any scalar m^k : different variables propagate the contraction at different rates and through the cross-equation dynamics. This is a statement about beliefs — relevant, for instance, when model forecasts are compared with survey expectations. The equilibrium itself involves only one-step expectations, and with heterogeneous attention the composed forecasts need not respect the model’s own cross-variable relations (§2.7).*

2.4 Beyond first order: risk

Do inattentive agents also under-react to future *uncertainty*? A linearized model cannot say, because certainty equivalence removes the risk term. The scalar example of §2.1 gave the answer for one variable; the next result gives it for a whole vector of forward variables.

Proposition 2 (Behavioral discounting of risk). *Hold the reference point $\bar{X}$ fixed. Let $G(\hat{x}) \equiv g(\bar{X} \odot e^{\hat{x}})$ be three times continuously differentiable, with third derivatives bounded by $Ce^{c\|\hat{x}\|}$ for some constants $C, c \geq 0$; power functions such as $g(X) = X^a$ qualify. Write $\hat{x}_{t+1} = \varsigma \tilde{x}_{t+1}$, with ς the perturbation scale, and let $\mathbb{E}_t[\|\tilde{x}_{t+1}\|^3 e^{\bar{c}\|\tilde{x}_{t+1}\|}] < \infty$ for some $\bar{c} > 0$; bounded third moments suffice when $c = 0$. Let $\mu \equiv \mathbb{E}_t \hat{x}_{t+1}$ and $\Sigma \equiv \text{Var}_t(\hat{x}_{t+1})$. To second order in ς ,*

$$\mathbb{E}_t^{\text{BR}} g(X_{t+1}) = g(\bar{X}) + \nabla^\top (D_\rho \mu) + \frac{1}{2} (D_\rho \mu)^\top H (D_\rho \mu) + \frac{1}{2} \text{tr}(H D_\rho \Sigma D_\rho) + o(\varsigma^2), \quad (12)$$

where ∇ and H are the gradient and Hessian of G at $\hat{x} = 0$. Hence the risk term loads the conditional second moments through

$$\frac{1}{2} \text{tr}(H D_\rho \Sigma D_\rho) = \frac{1}{2} \sum_{i,j} \rho_{X_i} \rho_{X_j} H_{ij} \Sigma_{ij} : \quad (13)$$

relative to rational expectations the loading on $\text{Cov}_t(\hat{x}_{i,t+1}, \hat{x}_{j,t+1})$ is scaled by $\rho_{X_i} \rho_{X_j}$, and each own-variance loading by $\rho_{X_i}^2$. With a common attention ρ the whole risk term is scaled by ρ^2 . With heterogeneous attention the attenuation is componentwise, and the sum in (13) need not shrink in magnitude.

With a fixed anchor and a common attention coefficient, then, the direct second-order loading on risk is multiplied by ρ^2 , and cognitively discounted agents under-react to future uncertainty — an effect that is identically zero in any linearized formulation. In the example

of §2.1, $G(c) = e^{-c}$ has $H = 1$ and the risk term is $\frac{1}{2}\rho^2 v$. With heterogeneous attention no aggregate statement follows, because unequal contraction can undo an offset between correlated exposures: for $G(z) = \frac{1}{2}(z_1 - z_2)^2$ and $z = (u, u)$ the rational risk term is zero, while $D_\rho = \text{diag}(1, \frac{1}{2})$ gives $\frac{1}{8}\text{Var}_t(u) > 0$. The comparison is at fixed (μ, Σ) : it isolates how the *loading* on given conditional moments changes with attention. In equilibrium Σ itself depends on ρ ; the identification analysis of §5.7.3 works with the full equilibrium mapping. When the reference point is random, the result applies to the deviations conditional on the anchor (§2.5). The proposition is written for log-chart variables; with mixed charts, $G(\hat{x}) \equiv g(\chi^{-1}(\chi(\bar{X}) + \hat{x}))$ and the proof is unchanged. The proof, with the remainder bound, is in Online Appendix OA-B.1.

The consumption Euler equation shows the economic content.

Corollary 1 (Precautionary wedge under cognitive discounting). *Consider the Euler equation $c_t^{-1} = R_t \beta \mathbb{E}_t^{\text{BR}} c_{t+1}^{-1}$ with c the cognitively discounted variable ($\rho \equiv \rho_c$) and a fixed reference point $\bar{c}$. If $\log c_{t+1}$ is conditionally Gaussian with mean deviation $\mu_{c,t} \equiv \mathbb{E}_t \hat{c}_{t+1}$ and variance $\sigma_{c,t}^2 \equiv \text{Var}_t(\hat{c}_{t+1})$, then, exactly,*

$$\log R_t = -\log \beta - \hat{c}_t + \rho \mu_{c,t} - \frac{1}{2} \rho^2 \sigma_{c,t}^2. \quad (14)$$

At $\rho = 1$ this is the standard lognormal Euler equation, $\log R_t = -\log \beta + \mathbb{E}_t \Delta \hat{c}_{t+1} - \frac{1}{2} \sigma_{c,t}^2$. The precautionary term $-\frac{1}{2} \rho^2 \sigma_{c,t}^2$, the fall in the required return that uncertainty about future consumption induces, is scaled by ρ^2 relative to its rational value. The conditional mean of the future log deviation is scaled by ρ , and current consumption is not contracted: $-\hat{c}_t + \rho \mu_{c,t} = \rho \mathbb{E}_t \Delta \hat{c}_{t+1} - (1 - \rho) \hat{c}_t$. Holding the conditional consumption distribution fixed, attention $\rho = 0.8$ gives the direct precautionary term in the required log return 64 percent of its rational-expectations magnitude. This is a statement about the required return, not about the quantity saved, which also depends on the budget constraint, prices and the equilibrium distribution of consumption. Without conditional Gaussianity the formula holds to second order rather than exactly; Online Appendix OA-B.2 gives the proof and a non-Gaussian example, and Online Appendix OA-B.5 verifies the ρ^2 scaling numerically.

Remark 4 (Discounting variables versus discounting terms). *A rule that discounts the expectation of a forward-looking term, $\mathbb{E}_t^{\text{BR}} g(X_{t+1}) = \bar{g} + M(\mathbb{E}_t g(X_{t+1}) - \bar{g})$, shrinks the expected deviation of the term by M , so the risk correction inside that expectation is weighted by M . With a common M and coherent reference values $\bar{g} = g(\bar{X})$ the rule is linear in g : it is the expectation under the mixture $(1 - M) \delta_{\bar{X}} + M P_t$ of the true conditional law P_t and a point mass at the reference. Splitting a term into additive pieces therefore does not change it; what must be chosen is which economic objects are discounted, and whether coefficients*

or reference values differ across them. The operator instead takes the expectation under the push-forward of P_t by the contraction of the variables inside g , so the second-order loading is weighted by ρ^2 and, by Proposition 3 below, the contribution of the variable's k -th one-step cumulant by ρ^k — the scalar example of §2.1. For a term with gradient a at the reference, the two rules have the same first-order loading if and only if $a^\top(D_\rho - MI) = 0$ on the relevant directions: matching common attention and reference values ($D_\rho = MI$) suffice, and a single term coefficient M cannot in general reproduce heterogeneous ρ_{X_i} . Their nonlinear risk loadings generally differ: with common attention the operator discounts risk more than the mean. Section 5.7.4 compares them in the estimated model.

2.5 Cumulants and finite beliefs

The mean action (7) and the variance action of Proposition 2 are the first two terms of a single statement about how the operator acts on the *whole* one-step conditional distribution of a future variable. In the level chart the statement is immediate: the map is affine, so it scales every k -th conditional cumulant of the deviation that exists by exactly ρ^k , with no convergence caveat. In the log chart it is a statement about the moment generating function, and it has a consequence: behavioral expectations can be finite where rational ones are not.

Proposition 3 (Cumulant contraction). *Let $g(X) = X^a$ (for instance $a = -\gamma$, the CRRA marginal utility), with X a scalar log-chart variable carrying attention ρ and a fixed reference point $\bar{X}$, and let $\mathcal{M}_t(s) \equiv \mathbb{E}_t e^{s\hat{x}_{t+1}}$ be the conditional moment generating function of the forward log deviation. Then, as an identity in $[0, \infty]$,*

$$\mathbb{E}_t^{\text{BR}} g(X_{t+1}) = \bar{X}^a \mathcal{M}_t(a\rho), \quad (15)$$

so the behavioral expectation is finite exactly when the moment generating function is finite at $a\rho$. If moreover $|a|\rho < r_t$, the radius of convergence of the cumulant series $\psi_t(s) \equiv \log \mathcal{M}_t(s) = \sum_{k \geq 1} \kappa_{k,t} s^k / k!$, with $\kappa_{k,t}$ the k -th conditional cumulant of $\hat{x}_{t+1}$, then

$$\mathbb{E}_t^{\text{BR}} g(X_{t+1}) = \bar{X}^a \exp\left(\sum_{k \geq 1} \kappa_{k,t} \frac{(a\rho)^k}{k!}\right), \quad (16)$$

so the contribution of the k -th conditional cumulant is scaled by ρ^k relative to rational expectations ($\rho = 1$): the mean by ρ , the variance by ρ^2 , and every higher unstandardized cumulant by the corresponding power of ρ . (The contraction changes scale, not shape: for $\rho > 0$ standardized skewness $\kappa_3/\kappa_2^{3/2}$ and excess kurtosis κ_4/κ_2^2 are unchanged.)

The two displays answer different questions. Equation (15) holds for every ρ and gives a

finite value wherever the moment generating function is finite at $a\rho$; the cumulant expansion (16) additionally requires evaluation inside the radius of convergence of the cumulant series, which can be smaller. The proof is in Online Appendix OA-B.3. The finiteness part has a direct corollary.

Corollary 2 (Finite behavioral expectations). *Let $\alpha_t^+ \equiv \sup\{s \geq 0 : \mathcal{M}_t(s) < \infty\}$ and $\alpha_t^- \equiv \inf\{s \leq 0 : \mathcal{M}_t(s) < \infty\}$ be the one-sided abscissas of convergence of the conditional moment generating function (so $r_t \leq \min\{\alpha_t^+, -\alpha_t^-\}$). Suppose the rational expectation diverges, $\mathbb{E}_t X_{t+1}^a = \infty$, but the moment generating function is finite on the relevant side of zero: $\alpha_t^+ > 0$ if $a > 0$, or $\alpha_t^- < 0$ if $a < 0$. Then for every attention*

$$\rho < \bar{\rho}_t \equiv \begin{cases} \alpha_t^+/a, & a > 0, \\ \alpha_t^-/a, & a < 0, \end{cases}$$

(a strictly positive threshold) the behavioral expectation $\mathbb{E}_t^{\text{BR}} g(X_{t+1})$ is finite; its cumulant series converges if, in addition, $|a|\rho < r_t$. If the relevant abscissa is zero, as for a Student- t deviation, no $\rho > 0$ suffices.

Cognitive discounting moves $a\rho$ inside the strip on which the moment generating function is finite; §4.2 works out an example with an exponential tail. The corollary is about one-step *beliefs*: it delivers finite perceived expectations, not stable solved dynamics, which are a property of the equilibrium (Remark 5). Its proof, and the case $\rho = \bar{\rho}_t$, are in Online Appendix OA-B.4.

The risk channel is a property of beliefs. The perceived conditional distribution of the chart deviation is the image of the true one under the linear map D_ρ (§2.7), so its conditional cumulant generating function is, exactly, $\tilde{\psi}_t(s) \equiv \log \tilde{\mathbb{E}}_t e^{s^\top \hat{x}_{t+1}} = \psi_t(D_\rho s)$ wherever the right side is finite. By multilinearity every joint cumulant of order k is scaled by the product of the attentions involved, $\tilde{\kappa}(\hat{x}_{i_1}, \dots, \hat{x}_{i_k}) = \rho_{i_1} \cdots \rho_{i_k} \kappa(\hat{x}_{i_1}, \dots, \hat{x}_{i_k})$. Proposition 3 evaluates this perceived generating function. Proposition 2 is its second-order truncation (the truncation concerns g , not beliefs), and Corollary 2 reflects that the perceived abscissas of convergence are the true ones divided by ρ . The ρ^k scaling is thus an exact property of what agents believe about next period, not an artifact of perturbation. Three qualifications keep the claim honest. (a) It concerns chart deviations: in the log chart the perceived *level* $\bar{X}e^{\rho\hat{x}}$ does not have its cumulants scaled by ρ^k . Nor does the scaling extend to the anchor when the reference point is random, as for the real quantities of the laboratory model, anchored at realized technology. Writing $\log X = A + U$ with A the log anchor and U the deviation,

$\log \tilde{X} = A + \rho U$ and $\text{Var}_t(\log \tilde{X}) = \text{Var}_t(A) + 2\rho \text{Cov}_t(A, U) + \rho^2 \text{Var}_t(U)$: anchor risk is not discounted, even at $\rho = 0$. With a random anchor, the results apply to the detrended deviations conditional on the anchor, and the Hessian in the expansion of Proposition 2 then depends on the anchor. The conditional results must be integrated over the anchor under the required integrability conditions. (b) It is one-step-ahead: composed k -step beliefs do not scale cumulants uniformly (Proposition OA-A.2(iv)). (c) It is an assumption, not a consequence of optimal attention: leading-order attention choice is indifferent between contracting the whole realization and contracting only its mean (Proposition OA-A.4(iv)). The pathwise contraction, and with it the discounting of risk, is a substantive behavioral hypothesis. It is also the hypothesis that separates the operator from other routes to a first-order discount (§2.7).

Remark 5 (The perturbation counterpart). *A related mechanism may operate on the perturbation solution of the nonlinear model; it is a possible mechanism that must be verified model by model, not a consequence of the results above. Proposition 3 concerns the cumulants of a given conditional distribution; a DSGE perturbation is an asymptotic series in the scale of fluctuations, and simulating an un-pruned higher-order approximation can generate explosive paths — the reason pruning schemes were introduced (Kim et al., 2008; Den Haan and De Wind, 2012; Andreasen et al., 2018; Lombardo and Uhlig, 2018). The behavioral operator contracts every forward deviation by ρ before the expansion is taken, so an order- k derivative of the equilibrium residual with respect to its forward arguments carries one factor of ρ per forward appearance — up to ρ^k for pure forward products, fewer for mixed terms. This does not translate into a rate for the solved coefficients: solving the equilibrium introduces inverse Jacobians, recursive products of lower-order coefficients, and a change in the equilibrium law itself. The growth rate of the behavioral coefficients is therefore not in general ρ times the rational one, and no universal attention threshold follows. (In the scalar equation $y_t = s + \eta_1 \mathbb{E}_t y_{t+1} + \eta_2 \mathbb{E}_t [y_{t+1}^2]$ the branch through zero has coefficients growing at rate $4\eta_2 \rho^2 / (1 - \eta_1 \rho)^2$; at $\eta_1 = \rho = \frac{1}{2}$ this is one ninth, not one half, of the rational rate.) Cognitive discounting is therefore a candidate structural, economically interpretable complement to pruning: it acts on divergence at its source, the forward feedback, rather than truncating it after the fact. Its effect on solved higher-order dynamics is a model-specific numerical question. Section 4.2 answers it for the laboratory model: in repeated simulations, low enough attention removes the explosions detected when divergence is driven by persistence (except at the highest persistence considered), but not, in general, explosions driven by large shocks.*

2.6 Trends and regimes

Models used for policy grow and switch regimes. Neither requires new machinery; both require a decision about what the contraction is anchored to.

Trends. A growing variable is contracted toward a reference point that moves with its trend. If $X_{t+1} = T_{t+1}x_{t+1}$, with T_{t+1} a trend known at $t+1$ (realized technology for the real quantities of the laboratory model) and x stationary, and the reference point is $\bar{X}_{t+1} = T_{t+1}\bar{x}$, then the perceived level is $\bar{X}_{t+1}(X_{t+1}/\bar{X}_{t+1})^\rho = T_{t+1}\bar{x}(x_{t+1}/\bar{x})^\rho$. Discounting the level is thus discounting its detrended counterpart, and the model in levels and its detrended counterpart have the same equilibrium conditions. What is not contracted is the trend itself, which serves as the anchor and is perceived correctly, and with it the risk in the trend (§2.5). Online Appendix [OA-C.1](#) gives the construction.

Regimes. With Markov regime switching three conventions define the operator: attention is indexed by the regime agents *currently* occupy; each continuation is anchored at the reference point of the regime being integrated over; and the transition probabilities are not misperceived.

Proposition 4 (Regime switching). *Let $r_t \in \{1, \dots, h\}$ follow a Markov chain with (possibly endogenous) transition probabilities $\pi_{r_t,j}(X_t)$, assumed measurable in X_t with $\sum_j \pi_{r_t,j} = 1$. The behavioral system*

$$0 = \mathbb{E}_t \sum_{j=1}^h \pi_{r_t,j}(X_t) f\left(\bar{X}(j) \odot (X_{t+1}(j) \oslash \bar{X}(j))^{\rho_X(r_t)}, X_t(r_t), X_{t-1}, \varepsilon_t, \theta(r_t)\right) \quad (17)$$

is a switching system of the same class as its rational counterpart (Maih, 2015): attention $\rho_X(r_t)$ is indexed by the regime agents currently occupy, while each continuation is anchored at the regime-appropriate reference point $\bar{X}(j)$ of the regime being integrated over. It nests:

- (i) the constant-parameter behavioral model (5) when ρ_X , $\bar{X}$, and θ are regime-invariant;*
- (ii) rational expectations regime-by-regime when $\rho_X(r_t) = \mathbf{1}$ for all r_t .*

Both nestings are identities on equations, so they hold exactly at every order of approximation. “Of the same class” is a statement about the form of the equations, not about their solutions: existence and determinacy of a solution of (17) are not implied and must be checked case by case (Online Appendix [OA-F.2](#) gives an example in which a solution that is stable regime by regime is found only below an attention threshold). The proof is in Online Appendix [OA-C.2](#).

Remark 6 (Attention that switches with the state). Equation (17) lets attention itself depend on the state of the economy: $\rho_X(r) < \rho_X(r')$ means agents discount the future more while the economy is in regime r than in r' . The leading application is the effective lower bound: if agents pay less attention to the future precisely while the policy rate is constrained, the model generates a regime-dependent expectations channel that a constant- ρ model cannot. This current-regime indexing of attention — with continuation anchors $\bar{X}(j)$ at the future regime — is as implemented in RISE (a correspondence we verified numerically). The online appendix uses it, at the calibration, with ρ_{normal} and ρ_{floor} set separately; the estimation keeps attention the same in and out of the bound.

Remark 7 (Exogenous versus endogenous attention). Throughout, ρ_X is a parameter, possibly regime-dependent. This is the right modeling choice if cognitive discounting is read as a deep cognitive trait. If instead ρ_X is read as attention in the sense of the rational-inattention and sparsity literatures (Sims, 2003; Maćkowiak and Wiederholt, 2009; Gabaix, 2014), it should respond to the stakes of forecasting. Whether those stakes rise with the conditional moments that the operator discounts depends on the decision problem: a higher variance need not raise the value of attention (the leading-order sparse-max loss of Proposition OA-A.4 depends only on conditional means), and signed higher cumulants do not order the stakes without restrictions on preferences, exposures, and the cost of attention. In an attention-choice model in which higher uncertainty does raise the marginal benefit of attention, agents would pay more attention in volatile states, raising ρ_X and partly undoing the fixed-attention contraction of risk. Gabaix (2016b) sketches such a case for the linear model: optimally chosen attention stays at its default for moderate volatility and rises once volatility or the stakes are large, and the survey evidence cited in the introduction is consistent with the same pattern. Two implications follow. First, in such a model, and holding the relevant distributions and exposures fixed, constant attention set at its tranquil-times value overstates the direct attenuation of risk loadings in volatile states, where the k -th cumulant would be scaled by $\rho_{X,t}^k$ with $\rho_{X,t}$ higher. Neither this monotone response of attention nor an equilibrium bound on the under-reaction to risk is proved here. Second, the operator accommodates endogenous attention without modification. The simplest route, already covered by Proposition 4, is to let attention switch with a regime whose transition probabilities depend on the state — for example on a volatility or inflation indicator — so that agents become more attentive when the economy moves into a high-uncertainty regime. A smooth alternative replaces the parameter by a function of predetermined states, $\rho_{X,t} = h(X_{t-1})$, which leaves the operator defined pointwise and, for smooth h and under the usual local regularity conditions, the system solvable by perturbation. Making attention depend on the conditional moments themselves, as a fully optimizing inattention model would, turns ρ into an equilibrium fixed

point, an extension not pursued here.

2.7 What the operator is

The operator was defined as a transformation of equilibrium conditions. What does it assume about beliefs, and how does it compare with other ways of discounting the future?

A rational expectation under a perceived law. Applying the operator to any equilibrium condition is the same as taking an ordinary expectation under a *perceived law* of next period's variables: the true conditional law of the behavioral equilibrium, with each selected variable's chart deviation contracted toward its anchor. What agents perceive correctly is the current state, the anchor state (in particular the future regime), and the regime transition probabilities. The identity is exact, at every order and for every function whose expectation is defined (any nonnegative or integrable f), and the perceived law is unique (Proposition OA-A.1). Two consequences matter for using the operator. First, the law that agents contract is the equilibrium law of the *behavioral* economy, not the rational-expectations law: the behavioral equilibrium is a fixed point in which the same transition generates the data and, contracted, the beliefs. Second, finiteness statements such as Corollary 2 are statements about that perceived law.

What the representation does not deliver. It represents the one-step beliefs that the operator imposes on the equilibrium conditions; it is not a derivation of the households' and firms' full dynamic problems. Read as a household's perceived problem it is local, because the contracted vector must stay in the admissible domain (§2.2).

Horizons. The equilibrium conditions involve only one-step-ahead expectations, so the representation is all the equilibrium requires. Multi-step forecasts are an additional object. Composing the one-step operator is exactly the same as iterating the perceived law, provided the transition is specified at the off-equilibrium points the perceived chain reaches. With a linear conditional mean, composed forecasts are exactly $(D_\rho A)^k$; beyond first order they generally are not. In general, no single map of this kind, independent of the current state, carries the family of true k -step laws onto the perceived ones for $k \geq 2$, so the ρ^k scaling of cumulants is a one-step property (Proposition OA-A.2).

Do perceived futures respect the model? Each forward variable is contracted by its own attention, so the perceived vector need not satisfy the equilibrium's own relations among variables. A linear relation among chart deviations, say $\hat{y} = \hat{c} + \hat{g}$ with $\hat{c}$ and $\hat{g}$ free

to vary, is respected by the perceived law only if $\rho_y = \rho_c = \rho_g$. A nonlinear relation, for example a resource identity in levels written in log charts, is not respected even under a common interior attention $\rho \in (0, 1)$ when it must hold on a neighborhood of the reference point (Proposition OA-A.3(b)). At $\rho = 1$ every relation is respected. This is where the operator departs from Gabaix’s construction. Gabaix (2020) contracts the *state*, with a perceived transition specified nonlinearly before it is linearized, and lets endogenous variables be equilibrium functions of the perceived state, so perceived futures are model-consistent by construction. The operator instead contracts the forward *variables* themselves. With common attention m on the relevant jump variables, the two agree on one-step beliefs at first order, which is all the equilibrium conditions use; composed forecasts can differ unless the state contraction is matched as well. Beyond first order the constructions can differ; where the equilibrium maps have nonzero quadratic curvature they generally differ already at second order (Corollary OA-A.1; Proposition OA-A.3(b)). In the three-equation application exogenous states never appear at $t + 1$, so their attention is neither pinned down by that model nor identified by its likelihood. (This is a property of the application, not of the operator: in the laboratory model of Section 3 the future preference shifter does appear at $t + 1$, and its attention is fixed at one by modeling choice.) The robust statement of the nesting is therefore that the operator reproduces Gabaix’s one-step beliefs and, through the forward solution of its one-step conditions, his m^k horizon discount.

A sparse-max reading. Gabaix’s (2014) sparse max (extended to dynamic problems in Gabaix, 2016a) has an agent first choose attention by minimizing a quadratic approximation of the loss from inattention plus a cost of attention, and then maximize exactly under perceived variables. With the default at the reference point and the perception rule applied to the realization of the forward variable, the second step is the operator exactly. Under stated approximations, the first step yields an attention of the form $\rho_i = \bar{\Lambda}_{ii}/(\bar{\Lambda}_{ii} + 2\kappa_i)$, with $\bar{\Lambda}$ the stakes and κ_i the weight of a quadratic attention cost $\kappa_i \rho_i^2$. Since the stakes depend on ρ , this is a fixed-point condition. The approximations are the local maximizer branch, a leading-order quadratic loss, diagonal stakes, and attention chosen once (Proposition OA-A.4 and Online Appendix §OA-A.5). And the leading-order loss depends on the perception rule only through its conditional mean, so attention choice cannot tell whole-realization contraction from mean-only contraction. The discounting of risk is therefore an additional behavioral hypothesis, not a consequence of optimal attention.

Other routes to a discount on the future. Several frameworks with quite different primitives deliver, in their linearized form, forward-looking equations whose expectational

terms carry a discount below one, and hence a reduced form close to Gabaix’s: incomplete information and higher-order uncertainty (Angeletos and Lian, 2018; Angeletos and Huo, 2021), level- k thinking and reflective equilibrium (Farhi and Werning, 2019; García-Schmidt and Woodford, 2019), finite planning horizons (Woodford, 2019), and uninsurable idiosyncratic risk with borrowing constraints (McKay et al., 2017). Suppose such a model maps into a first-order system of the form (10)–(11), with the same policy rule and the same shocks. The first-order results that depend only on that system then carry over, with ρ reinterpreted as that model’s effective discount: the relaxed Taylor principle and weaker forward guidance. A fiscal multiplier carries over only if the complete first-order system also matches, including how government spending and its financing enter the natural rate, the resource constraint and the policy closure; the two private equations do not determine it. The higher-order results do not carry over automatically. The discounting of risk (Proposition 2) and finite behavioral expectations (Corollary 2) follow from a particular assumption about the *perceived* distribution of the future: that it is the true one contracted toward the reference point, so that every (unstandardized) cumulant of the deviation is shrunk. Frameworks that generate the same first-order discount can imply different perceived higher moments. In the incomplete-markets route, for instance, the discount is itself a product of precautionary behavior against idiosyncratic risk, so that more risk strengthens the discount rather than being discounted by it. The frameworks are thus observationally close to first order, and their higher-order predictions can differ, which creates potential tests with suitable observables and identifying restrictions; this is where the operator makes a sharp prediction.

Departures that are not attenuation. Behavioral expectations need not take the form of attenuation at all. Under adaptive learning agents estimate the law of motion and update their estimates as data arrive (Evans and Honkapohja, 2001); Meggiorini and Milani (2026) compare learning with cognitive discounting directly. Diagnostic expectations *over-react* to news (Bordalo et al., 2018; L’Huillier et al., 2024), and in behavioral learning equilibria agents forecast with simple autoregressive rules disciplined by observed autocorrelations (Hommes et al., 2023). The operator is not designed to represent these schemes. Formally, a value $\rho_X > 1$ would extrapolate deviations from the reference point and is a crude caricature of over-reaction; by Proposition 3 it would *amplify* every cumulant by ρ^k , so that the finiteness result is reversed. The stabilizing properties established here are therefore specific to attenuation, not general properties of behavioral expectations.

2.8 Using the operator

For a researcher who wants to replace rational expectations in an existing nonlinear model, the preceding results reduce to six steps.

1. *Choose coordinates and anchors.* Give each forward variable a chart (log for positive levels, identity for rates and gaps) and a reference point: the steady state, a regime-specific value, or a point that moves with the trend. Decide whether the anchor is fixed or random; risk in a random anchor is not discounted (§2.5). Check that the contracted vector stays in the admissible domain of the equilibrium conditions (§2.2).
2. *Choose the contracted variables and their attention.* Select the forward variables to discount and give each an attention: common or variable-specific, constant, regime-dependent, or a function of predetermined states. Composed forecasts compound as m^k when a constant common discount also covers the exogenous states, at first order or exactly for a linear conditional-mean transition (Proposition 1); this forecast property is distinct from the m^k discount that forward iteration of the linear one-step equilibrium conditions produces.
3. *Apply the same contraction at every selected occurrence.* Every occurrence of a selected variable dated $t + 1$, in every equation, is replaced by its contraction (4); no term-by-term choice remains. This substitution, with attention a (possibly regime-switching) parameter, is as implemented in RISE, so the object solved is the object analyzed.
4. *Solve, and check existence and accuracy.* The transformed model is of the same class as the rational one, but existence, determinacy and stability are not inherited. Check them: the root count at first order (regime by regime is not the same as mean-square stability in a switching model), the behavior of simulated paths at higher order (pruned and un-pruned), and the accuracy of the approximation.
5. *Recover the rational limit.* At $\rho_X = 1$ the operator is the identity on equations, so the rational model must be reproduced exactly; for a trending model the levels model must also reproduce its detrended twin.
6. *Separate direct belief effects from equilibrium effects.* Propositions 2–3 hold the conditional distribution fixed; in equilibrium attention also changes that distribution. Comparisons across attention levels should say which of the two they report.

Two questions remain application-specific: whether attention is identified separately from other features of the model, and whether restrictions such as a common or block-specific attention are supported. The paper’s own answers, for the laboratory model, are in Section 5.

Remark 8 (Identification). *Estimation must separate two forces that both compress the conditional distribution of future variables in a given regime: a low $\rho_X(r)$ (inattention) and a low conditional shock volatility $\sigma(r)$. The risk channel of Proposition 2 can in principle help separate them: inattention scales the risk loading by $\rho_X(r)^2$ while leaving the realized shock variance untouched, whereas a change in $\sigma(r)$ scales the variance itself. The two therefore leave different signatures on the higher-order dynamics (e.g. the curvature of impulse responses and the order-dependence of moments), which a nonlinear likelihood can in principle exploit. Whether the confound is present in a given application, and whether the nonlinear likelihood is needed to resolve it, is an empirical question; §5.6 and §5.7.3 answer it for the estimated model, where the local diagnostics show no severe confound at first order. Restricted attentions (sector-specific, block-specific, or one for all variables), compared by marginal likelihood, provide the parsimonious fallback when the data are uninformative about variable-specific attention.*

3 The quantitative laboratory

The quantitative results use a single model, chosen so that each property of the operator established in Section 2 can be put to work and checked in it. The optimal-policy exercises use a stationary variant of it, described in §4.7. This behavioral New Keynesian model — henceforth the *laboratory model* — is a medium-scale nonlinear economy with four features, each of which corresponds to one of those properties. Its equilibrium conditions are listed in full in Appendix A.

- *Nonstationary.* The model has a balanced-growth path: technology, output, consumption, the real wage, and prices are nonstationary in logs. The operator’s reference point $\bar{X}$ moves with the balanced-growth path, shifted by realized technology and the lagged price level, so the contraction in (6) is taken toward a moving anchor (§2.6; Appendix A).
- *Nonlinear.* Internal habit formation, Rotemberg price *and* wage adjustment with indexation, and a credit-spread block make the equilibrium conditions genuinely nonlinear.
- *Multiple forward terms per equation.* The wage and price Phillips curves each contain several forward-looking terms ($\mathbb{E}_t \lambda_{t+1}, \mathbb{E}_t p_{t+1}, \mathbb{E}_t y_{t+1}, \mathbb{E}_t \pi_{t+1}^w$) — the exact configuration that Gap 2 says the basic device cannot handle. The model motivates the operator rather than merely hosting it.

- *An occasionally binding lower bound.* The policy rate is $R_t = \max(\underline{R}, R_t^{\text{Taylor}})$: a hard kink that binds only when the shadow Taylor rate falls below the floor, so whether and how long the bound binds is an equilibrium outcome. Attention is the same in and out of the bound. (A variant in which the bound is a Markov regime in the style of [Binning and Maih \(2017\)](#), and attention may differ at the floor, is studied in the online appendix.)

The model is solved by perturbation up to order 3 in the exercises below (higher orders are covered by the analytical example). Lower-bound episodes are computed at first order under perfect foresight, with the anticipated-shocks approach of [Maih \(2010\)](#). The bound is enforced by shocks that enter only the policy-rate equation and are separate from the monetary policy shock. Their path makes the complementarity conditions of the max operator hold: they are nonnegative, and nonzero only in quarters where the bound binds. One check validates the implementation: the rational-expectations limit is exact by construction. Setting $\rho_X = \mathbf{1}$ leaves every equilibrium condition unchanged, and at $\rho_X = \mathbf{1}$ the levels model reproduces its detrended twin to machine precision. Section 5 estimates the model on seven observables over a pre-COVID sample. Throughout, the three-equation New Keynesian model of [Gabaix \(2020\)](#) carries the pencil-and-paper analytics (Proposition 1 and the determinacy region below), while the laboratory model carries the quantitative claims.

4 Results: beyond linear cognitive discounting

This section reports seven findings. Where an exercise repeats one in [Gabaix \(2020\)](#), we compare our results with his; where it has no linear counterpart, we say so and contrast it with the linearized model.

One attention, for tractability. Each exercise in this section uses a single attention ρ , common to all discounted variables. Three parameter points are used. The first is the calibration. The second is the *restricted estimate*: the posterior mode of the common-attention restriction of §5.7, with $\hat{\rho} = 0.897$, reported alongside the posterior mode of the rational-expectations model. The third, labeled “restr.” in tables and figures, is the calibration with attention set to $\hat{\rho}$. With attention estimated variable by variable, as in the baseline estimation of Section 5, the results depend on the whole attention vector, which this section does not evaluate.

4.1 Cognitive discounting relaxes the Taylor principle

The clearest analytical payoff of Proposition 1 concerns determinacy. In the discounted three-equation model, the region of the parameter space (ϕ_π, m) that delivers a unique stable equilibrium expands as attention falls (Figure 1). Under rational expectations ($m = 1$) the determinacy threshold is the textbook Taylor principle, $\phi_\pi^* = 1$; it falls with m and reaches $\phi_\pi^* = 0$ at $m \approx 0.73$. Cognitive discounting relaxes the Taylor principle: myopia is itself stabilizing, because agents who discount the future put less weight on the self-fulfilling expectations that would otherwise require an aggressive policy response (Gabaix, 2020).

Since this repeats an exercise of Gabaix (2020), the comparison should be exact. Writing m also for his attention (his M) and setting his firm coefficient to $M^f = m$ (Remark 2), his criterion (his Proposition 3.1) and the operator’s differ only on the right-hand side. Because the operator also discounts expected inflation in the real rate, its criterion,

$$\phi_\pi + \frac{1 - \beta m}{\kappa} \phi_x + \frac{(1 - \beta m)(1 - m)}{\kappa \sigma} > m,$$

has m where Gabaix’s has 1, and is easier to satisfy (Online Appendix OA-D.1 derives both). Under a permanent interest-rate peg only the last term remains. At his calibration ($\kappa = 0.11$, $\sigma = 0.2$) it exceeds one for $m \lesssim 0.856$, so at his $m = 0.85$ the economy is determinate *even under a permanent peg*. Our calibration has a similar Phillips-curve slope ($\kappa = 0.10$) but a much larger intertemporal elasticity ($\sigma = 1$), so the peg becomes determinate only for $m \lesssim 0.733$. The direction is identical. The numerical threshold depends mainly on σ , and partly on the operator: at our calibration, Gabaix’s IS curve would give 0.687 instead of 0.733. (The Online Appendix also reports the cutoff under Gabaix’s originally reported, compounded firm coefficient.)

The same logic delivers a sharp threshold in the laboratory model, where the relevant policy experiment is a permanent peg at the lower bound. With the policy rate pegged, the Taylor feedback that otherwise anchors expectations is absent, and determinacy must come from the expectational friction alone — the role played by frictions in social memory and coordination in Angeletos and Lian (2023). At the kinked bound we use the eigenvalue criterion for the pegged system. It is a local test for a permanent peg, not a proof of global existence or uniqueness for the occasionally binding system (see Holden, 2023 for that question). The pegged system has a unique bounded solution only when attention is below a threshold ρ^* : $\rho^* = 0.736$ at the calibration and 0.720 at the restricted estimate. Under rational expectations, and for any ρ above the threshold, a permanent peg is indeterminate by this test. The threshold is robust to how the bound is modeled: when the floor is a Markov regime with a stochastic exit, the floor regime’s forward eigenvalue reaches one at

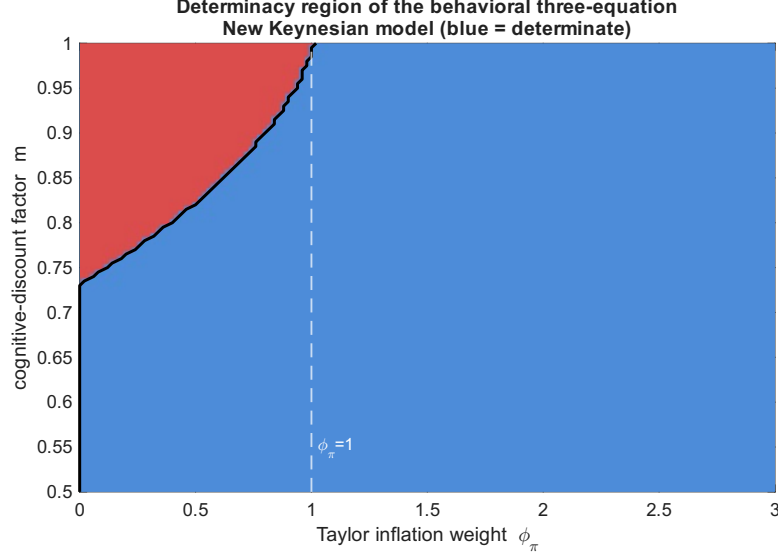


Figure 1: Determinacy region of the discounted three-equation New Keynesian model in (ϕ_π, m) . Cognitive discounting (lower m) relaxes the Taylor principle: the determinacy threshold ϕ_π^* falls from 1 at $m = 1$ to 0 at $m \approx 0.73$.

about $\rho = 0.732$, and above it the solution algorithms we use find no solution that is stable regime by regime (online appendix, Section OA-F.2). Section 4.4 shows that this threshold governs calendar forward guidance at the bound, and that the common attention of the restricted estimate lies above it; with attention estimated variable by variable, the question depends on the whole attention vector.

4.2 Higher-order dynamics: what attention does and does not stabilize

Section 2.5 shows that the operator contracts every forward-looking deviation, and with it the higher-order terms that load on products of forward variables (Remark 5). Whether this matters in a quantitative model depends on whether those terms are what makes a higher-order solution misbehave. We therefore ask two questions: does cognitive discounting change the higher-order solution at the baseline calibration, and does it remove explosive simulated paths where the rational solution diverges?

At the baseline calibration, it does not need to. Without the lower bound, the laboratory model's perturbation solution is well behaved under both expectation schemes. The peak absolute response of output to a monetary-policy shock is between 6.7×10^{-4} and 6.8×10^{-4} under rational expectations and between 4.7×10^{-4} and 4.8×10^{-4} under cognitive discounting ($\rho = 0.8$) at orders 1, 2, and 3, pruned or not (Table 1). Attention scales the response down

by about 30 percent; it does not rescue anything, because there is nothing to rescue.

Table 1: Peak $|\text{IRF of } Y|$ to a monetary shock, by perturbation order, without the lower bound, at the calibration. Under both schemes the responses are finite and of similar size at every order, with or without pruning.

peak $ \text{IRF of } Y $, no lower bound	order 1	order 2	order 3
RE (un-pruned)	6.8e-04	6.8e-04	6.7e-04
RE (pruned)	6.8e-04	6.8e-04	6.7e-04
CD ($\rho = 0.8$, un-pruned)	4.8e-04	4.7e-04	4.8e-04
CD ($\rho = 0.8$, pruned)	4.8e-04	4.8e-04	4.8e-04

Where the rational solution diverges, attention removes one kind of explosion but not the other. To find out, we push the model toward explosiveness along two directions and record, in repeated un-pruned simulations, whether paths explode by the criterion of Table 2 (five seeded simulations of 2,000 quarters; inflation non-finite or its standard deviation above 100 times its first-order value; Figure 2). Passing this criterion is evidence of no detected explosion, not a proof that all paths are bounded or that a stationary distribution exists. A simulation that stopped with a numerical error was also counted as explosive, and the stored results do not separate the two cases; this can only raise the counts, so no entry reported as bounded is affected. The first direction scales up the standard deviations of all structural shocks; the second raises the persistence of the five autoregressive shocks toward one. Pruned rational simulations show no explosion anywhere on both directions, as pruning is designed to ensure. The un-pruned rational solution explodes along both, and the operator behaves very differently on the two.

- *Persistence-driven explosions are removed.* At third order, un-pruned rational paths explode in four of five simulations at shock persistence 0.99 and in every simulation beyond. Under cognitive discounting no explosion is detected for $\rho \leq 0.95$ at persistence 0.99, for $\rho \leq 0.90$ at 0.995, and for $\rho \leq 0.70$ at 0.998; only at 0.999 does no ρ on the grid suffice. The more nearly the forward variables inherit a unit root, the lower attention must be — consistent with the mechanism in Remark 5. (At order 2, near-unit-root persistence also produces explosions at some intermediate values of ρ .)
- *Scale-driven explosions are not.* Once the shocks are large enough for most un-pruned rational paths to explode (scale 70 at order 2, 30 at order 3), no $\rho < 1$ on the grid removes the explosions. At somewhat smaller scales, some intermediate values of ρ explode at order 2 where rational expectations do not.

The split is consistent with the mechanism of Remark 5. The operator contracts forward terms toward their reference: it tames the propagation of expected future deviations, which

is what persistence amplifies, but it does not shrink the size of the shocks entering the quadratic and cubic terms today. In this model, cognitive discounting therefore acts as a *structural* stabilizer for one well-defined source of higher-order divergence, not as a substitute for pruning in general; the two are complements. (Pruning, for its part, removes a particular feedback in the approximation; it is not an unconditional guarantee of bounded paths for arbitrary shocks.) There is no counterpart to this exercise in Gabaix (2020): his model is linear, a first-order solution is certainty-equivalent, and the divergence studied here never arises.

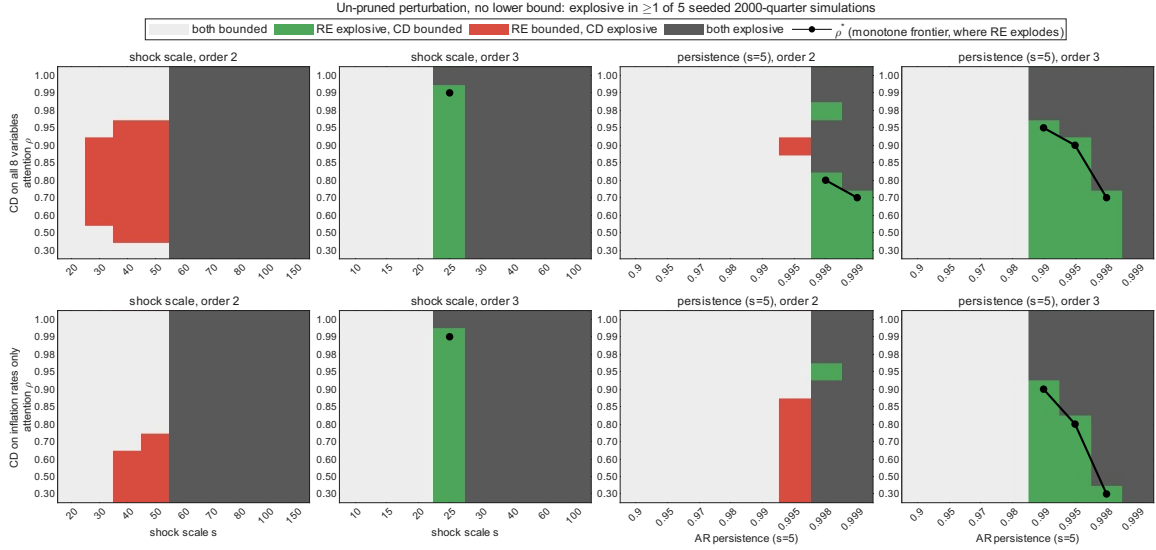


Figure 2: Stability frontier of the un-pruned perturbation solution; the same experiments as Table 2. When divergence is driven by persistence, cognitive discounting yields no detected explosion for $\rho \leq 0.95$ at persistence 0.99 and for $\rho \leq 0.70$ at persistence 0.998. At persistence 0.999, no grid value suffices. At order 2 some intermediate values of ρ explode where rational expectations do not, and explosions driven by large shocks are not removed in general.

A closed-form example shows the one-step mechanism of Corollary 2. Take a single forward variable with reference point normalized to $\bar{X} = 1$ (so $\hat{x} = \log X$), $\log X \sim \text{Exp}(\lambda)$ and $g(X) = X^a$, $a > \lambda$, so the rational expectation $\mathbb{E}_t X_{t+1}^a$ is *infinite*. By Proposition 3 the behavioral expectation is $\lambda/(\lambda - a\rho)$, finite for every $\rho < \lambda/a$: cognitive discounting brings a divergent rational expectation into the convergent region. Figure 3 and Table 3 show the partial sums $S_K(\rho) = \sum_{k \leq K} (a\rho/\lambda)^k/k$ of the cumulant (log moment generating function) series, whose limit S_∞ is the log of the behavioral expectation. Under rational expectations ($\rho = 1$) the sums explode with the truncation order K , with $S_{40} \approx 8.8 \times 10^5$. For $\rho < \lambda/a$ they converge. Here the divergence comes from the tail of a one-step distribution, not from persistence; the example illustrates the corollary, and the laboratory model shows which dynamic divergences the mechanism reaches.

Table 2: Stability frontier of the un-pruned perturbation solution (no lower bound). “RE: explosive paths” counts, out of five seeded simulations of 2,000 quarters, those in which inflation becomes non-finite or its standard deviation exceeds 100 times its first-order value. “CD: bounded for $\rho \in$ ” lists the values of $\rho < 1$ on the grid $\{0.99, 0.98, 0.95, 0.9, 0.85, 0.8, 0.7, 0.6, 0.5, 0.3\}$ at which none of the five simulations explodes by this criterion, under two discounting specifications: all eight forward-looking variables discounted (the paper’s specification), or only the four inflation rates. “Bounded” means no detected explosion, not a proof of boundedness. A simulation stopped by a numerical error counts as explosive; the stored results do not separate the two. Shock scale s multiplies all structural shock standard deviations (persistence 0.85); persistence sets the five autoregressive coefficients (shocks scaled by $s = 5$).

direction	value	order	RE: explosive paths		CD: bounded for $\rho \in$	
			un-pruned	pruned	all 8 variables	inflation rates
shock scale s	20	2	0/5	0/5	all	all
shock scale s	30	2	0/5	0/5	0.99–0.95; 0.50–0.30	all
shock scale s	40	2	0/5	0/5	0.99–0.98; 0.30	0.99–0.70
shock scale s	50	2	0/5	0/5	0.99–0.98; 0.30	0.99–0.80
shock scale s	60	2	1/5	0/5	none	none
shock scale s	70	2	4/5	0/5	none	none
shock scale s	80	2	5/5	0/5	none	none
shock scale s	100	2	5/5	0/5	none	none
shock scale s	150	2	5/5	0/5	none	none
shock scale s	10	3	0/5	0/5	all	all
shock scale s	15	3	0/5	0/5	all	all
shock scale s	20	3	0/5	0/5	all	all
shock scale s	25	3	1/5	0/5	all	all
shock scale s	30	3	5/5	0/5	none	none
shock scale s	40	3	5/5	0/5	none	none
shock scale s	60	3	5/5	0/5	none	none
shock scale s	100	3	5/5	0/5	none	none
persistence	0.9	2	0/5	0/5	all	all
persistence	0.95	2	0/5	0/5	all	all
persistence	0.97	2	0/5	0/5	all	all
persistence	0.98	2	0/5	0/5	all	all
persistence	0.99	2	0/5	0/5	all	all
persistence	0.995	2	0/5	0/5	0.99–0.95; 0.85–0.30	0.99–0.90
persistence	0.998	2	1/5	0/5	0.98; 0.80–0.30	0.95
persistence	0.999	2	3/5	0/5	0.70–0.30	none
persistence	0.9	3	0/5	0/5	all	all
persistence	0.95	3	0/5	0/5	all	all
persistence	0.97	3	0/5	0/5	all	all
persistence	0.98	3	0/5	0/5	all	all
persistence	0.99	3	4/5	0/5	0.95–0.30	0.90–0.30
persistence	0.995	3	5/5	0/5	0.90–0.30	0.80–0.30
persistence	0.998	3	5/5	0/5	0.70–0.30	0.30
persistence	0.999	3	5/5	0/5	none	none

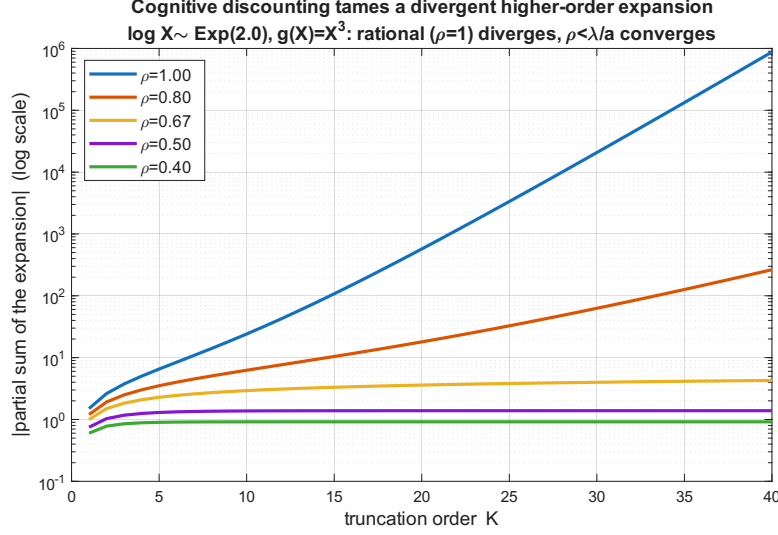


Figure 3: Corollary 2 in closed form. Partial sums $S_K(\rho) = \sum_{k \leq K} (a\rho/\lambda)^k/k$ of the log moment generating function (cumulant) series for $\log X \sim \text{Exp}(\lambda)$, $\bar{X} = 1$, $g(X) = X^a$ with $a > \lambda$; the behavioral expectation is $\exp(S_\infty)$. Under rational expectations ($\rho = 1$) the series diverges; cognitive discounting ($\rho < \lambda/a$) converges (Corollary 2).

Table 3: Partial sums S_K of the log moment generating function series by truncation order K : divergence under rational expectations, convergence under cognitive discounting for $\rho < \lambda/a$.

partial sum $S_K(\rho)$	$K=5$	$K=10$	$K=20$	$K=40$
$\rho = 1.00$	6.53	24.1	572	8.77e+05
$\rho = 0.80$	3.51	6.25	18	264
$\rho = \lambda/a = 0.67$	2.28	2.93	3.6	4.28
$\rho = 0.40$	0.9	0.916	0.916	0.916

4.3 Impulse responses and business-cycle volatility

How does cognitive discounting change the way the economy responds to shocks? At first order it dampens and reshapes the impulse responses relative to rational expectations (Figure 4, left): the peak output response to a monetary shock falls from 6.8×10^{-4} to 4.8×10^{-4} at $\rho = 0.8$, and monotonically with ρ (Figure 4, right). For a shock of this size the third-order responses are within 2% of the first-order ones at every ρ : attention, not the higher-order terms, drives the difference. The higher-order terms matter where the model is pushed toward divergence (§4.2).

The effect on volatility is selective (Table 4). Cognitive discounting ($\rho = 0.8$) lowers the standard deviation of inflation to 86% of its rational-expectations value but leaves that of output growth essentially unchanged. Cognitive discounting should not be summarized as simply making the economy “less volatile”.

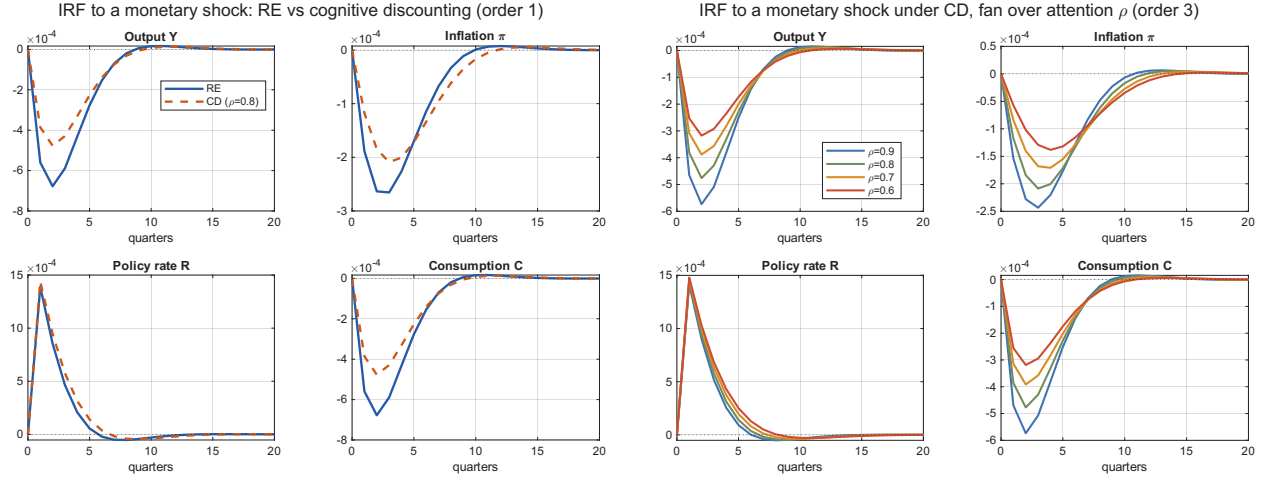


Figure 4: Left: order-1 impulse responses, rational expectations versus cognitive discounting ($\rho = 0.8$). Right: order-3 impulse responses as a fan over attention $\rho \in \{0.9, 0.8, 0.7, 0.6\}$. Monetary-policy shock, no lower bound.

Table 4: Business-cycle volatility without the lower bound, order 1: cognitive discounting versus rational expectations.

standard deviation, no lower bound	inflation	output growth
RE	0.0024	0.0045
CD ($\rho = 0.8$)	0.0021	0.0045
ratio CD/RE	0.857	1.005

4.4 Forward guidance: the puzzle, and how attention resolves it

The sharpest test of a behavioral expectations device is the *forward-guidance puzzle*: in the standard New Keynesian model an interest-rate change promised for the far future has an implausibly large effect on output and inflation today — an effect that, as the horizon lengthens, fails to die out and in fact *grows* (McKay et al., 2016; Del Negro et al., 2023). We run exactly that experiment, in the analytical three-equation behavioral model (the vehicle of Proposition 1) with the policy rate held exogenous: announce a one-period 25bp cut T quarters ahead and read the response of output and inflation *today*, against the horizon T (Figure 5, Table 5). Stacking $z = [x, \pi]'$, the system is $z_t = F \mathbb{E}_t z_{t+1} + G i_t$, so today's response to a cut at horizon T is $F^T G$, governed by the dominant eigenvalue of F (Online Appendix OA-D.2 gives F and G).

Under rational expectations that eigenvalue is $1.36 > 1$, so the effect *grows* with the horizon: a cut announced twenty quarters out moves output today by roughly 290 times the impact of an immediate cut of the same size. This is the puzzle in its starkest form. Cognitive discounting contracts F : the eigenvalue falls below one once attention $m \lesssim 0.73$, and the today-response then *decays* with the horizon, the mechanism Gabaix (2020) identified. At $m = 0.70$ the eigenvalue is 0.96 and the twenty-quarter effect is already below the contemporaneous one; at $m = 0.50$ it is 0.68 and guidance about the distant future is all but powerless. The operator therefore *nects* Gabaix's resolution of the forward-guidance puzzle and pins down its mechanism: the puzzle is a property of the dominant forward eigenvalue, which attention scales toward zero. The resolution is not automatic, though — near-rational attention does little (at $m = 0.99$ the eigenvalue is still 1.35 and the puzzle essentially survives); in this frictionless model it takes substantial myopia. Gabaix (2020) runs the same thought experiment (his Figure 2, after McKay et al., 2016) but cuts the *real* rate: the output response today is then σm^T , which dies out with the horizon for any $m < 1$, whatever the Phillips-curve slope. With the *nominal* rate held fixed, as here, inflation feeds back into the real rate and the forward eigenvalue depends on κ and σ ; at his calibration it is below one only for $m \lesssim 0.856$. As with determinacy, the threshold ($m \lesssim 0.73$ here) is a property of that structure, mainly of the intertemporal elasticity, and partly of the operator (§4.1); the two thresholds are the same object, since a pegged rate leaves only the forward matrix. What supplies the missing discipline in a realistic model is *structure*, to which we now turn.

The same lesson inside the quantitative model. In the laboratory model the forward channel is summarized by the spectral radius of the forward-propagation matrix $\mathcal{F}$: the impact today of a shock anticipated k quarters ahead is $\mathcal{F}^k R_0$ (cf. McKay et al., 2016).

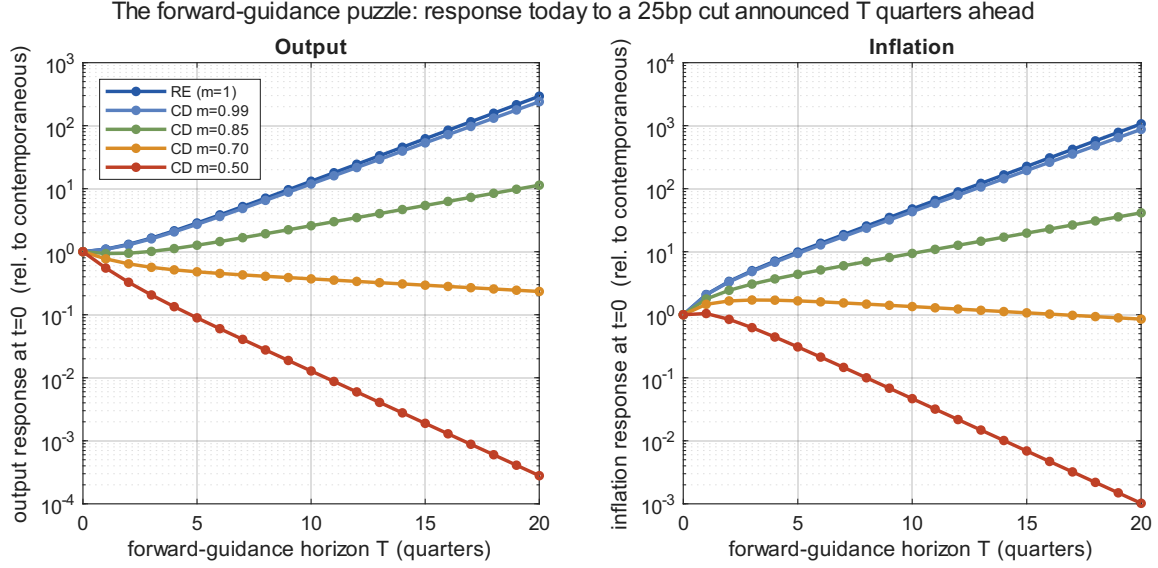


Figure 5: The forward-guidance puzzle in the three-equation behavioral model. Response of output (left) and inflation (right) *today* to a 25bp policy-rate cut announced T quarters ahead, relative to a contemporaneous cut (log scale), by attention m . Under rational expectations ($m = 1$) the effect grows with the horizon — the puzzle; cognitive discounting bounds and then reverses it once m is small enough that the dominant forward eigenvalue drops below one.

Table 5: Forward-guidance puzzle: dominant forward eigenvalue $\max|\lambda(F)|$ and the output response today to a cut T quarters ahead (relative to a contemporaneous cut), by attention m .

output response at $t=0$ (rel. to $T=0$)	$\max \lambda(F) $	$T=0$	$T=4$	$T=8$	$T=12$	$T=20$
RE ($m=1$)	1.364	1	2.15	7.07	24.4	293
CD $m=0.99$	1.351	1	2.06	6.52	21.6	240
CD $m=0.85$	1.160	1	1.12	1.93	3.47	11.4
CD $m=0.70$	0.955	1	0.515	0.408	0.338	0.234
CD $m=0.50$	0.682	1	0.134	0.0276	0.00596	0.000279

The radius is a system-wide diagnostic of long-horizon amplification. Below one, every anticipated-shock response eventually decays, though not necessarily monotonically. At or above one, the puzzle arises for a given shock and variable only if the shock excites a non-decaying mode. That mode need not be the dominant one, but it must show in the reported variable and not be canceled in the response. These conditions must be checked separately. Away from the bound, at the calibration, that eigenvalue is already 0.93 under rational expectations, consistent with habit and policy smoothing supplying the structural discipline the frictionless model lacked. Lowering attention with all other parameters fixed drives it further down, to 0.87 at $\rho = 0.897$ and 0.65 at $\rho = 0.6$ (Figure 6, Table 6), so announced guidance decays faster at long horizons. Across the two posterior modes (rational expectations and the common-attention restriction) the comparison reverses slightly: the eigenvalue is 0.913 at the rational mode and 0.917 at the restricted mode, because the other estimated parameters differ. Holding those at the restricted mode, moving ρ from 1 to 0.897 lowers it from 0.963 to 0.917.

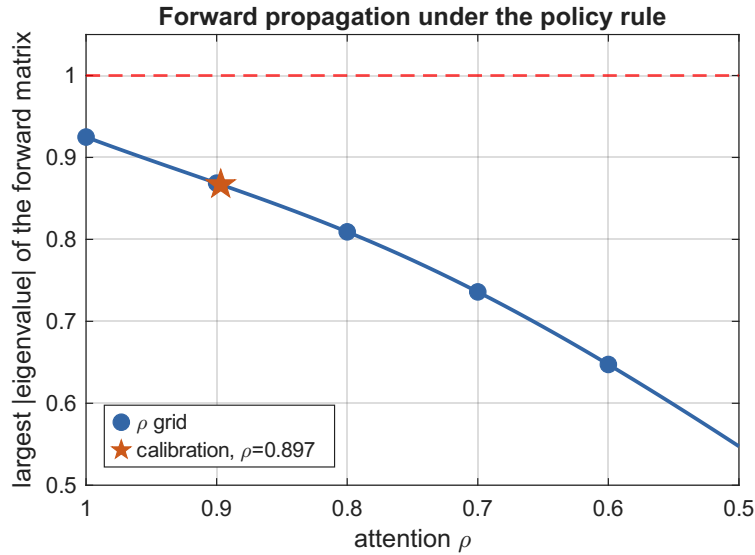


Figure 6: Forward-guidance propagation away from the lower bound. The spectral radius of the forward-propagation matrix falls from 0.93 (rational expectations) as attention ρ falls, other parameters at the calibration. The point labeled “restricted estimate” is the common attention of the restricted estimate, $\rho = 0.897$ (§5.7).

At the lower bound the puzzle returns, and a common attention of 0.897 does not remove it. The diagnostic above keeps the Taylor rule active; a lower-bound episode is precisely when that feedback is gone. We therefore run the calendar experiment where it matters. Start from a demand shock that holds the economy at the bound for eight quarters, and let the central bank commit, at the outset, to keep the rate at the floor for K quarters

Table 6: Maximum forward-propagation eigenvalue versus attention ρ , away from the lower bound, other parameters at the calibration. “restr.”: the common attention of the restricted estimate (§5.7). The last row reports only whether the spectral radius is at least one; by itself this does not establish a puzzle for a particular shock and variable, which also requires the shock to excite, and the variable to show, a non-decaying mode.

ρ	1.0	0.9	0.8	0.7	0.6	0.897 (restr.)
$\max \lambda(\mathcal{F}) $	0.925	0.869	0.809	0.736	0.647	0.867
$\max \lambda(\mathcal{F}) \geq 1?$	no	no	no	no	no	no

beyond the natural lift-off. Table 7 and Figure 7 report the effect on output today, relative to $K = 1$.

Under rational expectations the effect grows with K over the computed horizons — geometrically, by about 25% per additional quarter on average over the first six years: the forward-guidance puzzle in its starkest form, consistent with the indeterminacy of the permanent peg. (Each finite calendar peg, with its return to the Taylor rule, is a well-defined experiment; the growth over the computed range is a numerical observation, not a statement about the limit.) Cognitive discounting with a common attention of 0.897 does *not* remove it. At the restricted estimate ($\hat{\rho} = 0.897$) the effect grows geometrically with the peg length, at the same average rate as under rational expectations; at the calibration it still grows at $\rho = 0.8$; and at all these values a permanent peg is indeterminate by the local test (the pegged system has no unique bounded solution). For some peg lengths no equilibrium keeps the rate at the floor over exactly the forced window, and the table marks these cells “n.e.”. At the calibration, this occurs for $10 \leq K \leq 23$ at $\rho = 0.9$ and for $11 \leq K \leq 23$ at $\rho = 0.897$ (displayed: $K = 12\text{--}20$). At the restricted estimate, it occurs for $19 \leq K \leq 24$ at $\rho = 0.8$ (displayed: $K = 20, 24$). Every displayed entry is the solution whose binding set is exactly the forced window, a single linear system with one solution when it is nonsingular. A broader, non-exhaustive search that also allows an extra binding block before the episode or after the window finds additional candidate patterns in some rows. They keep the rate at the floor outside the window, so they are different policy paths, not further solutions of the forced-window problem. Only at $\rho = 0.6$ does the effect converge (at the calibration to about 76 times the one-quarter effect), and the permanent peg becomes determinate.

The dividing line is the peg-determinacy threshold of §4.1: a permanent peg has a unique bounded solution only for $\rho < \rho^*$, with $\rho^* = 0.736$ at the calibration and 0.720 at the restricted estimate. When the bound is instead modeled as a Markov regime, the floor regime’s forward eigenvalue crosses one at a close attention, about 0.732 (online appendix, Section OA-F.2). The operator thus nests Gabaix’s resolution of the puzzle: below ρ^* calendar guidance is bounded, and the more so the lower is ρ . But at a common attention

of 0.897, the restricted estimate, a permanent peg at the bound is not determinate, and the effect of finite calendar pegs keeps growing with their length over the horizons computed. With attention estimated variable by variable, determinacy depends on the whole attention vector, which we do not evaluate here. The threshold ρ^* describes a path on which all discounted variables share one attention; comparing individual attentions with it has no implication for determinacy.

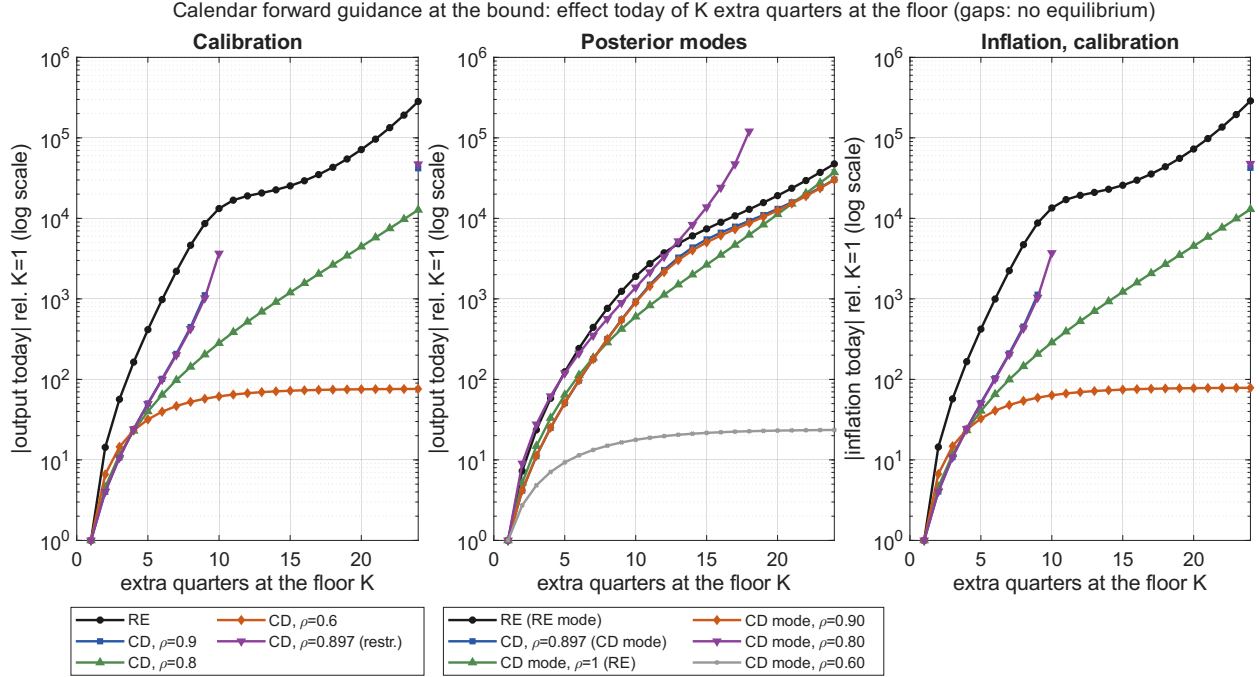


Figure 7: Calendar forward guidance at the lower bound. Effect on output today (left: calibration; center: posterior modes, with the restricted estimate for cognitive discounting) and on inflation today (right: calibration) of a commitment to keep the rate at the floor for K quarters beyond lift-off, relative to $K = 1$, log scale. The effect grows with K under rational expectations and at the common attention of the restricted estimate, $\rho = 0.897$, and converges only for attention below $\rho^* \in [0.720, 0.736]$ ($\rho = 0.6$). For some attentions above the threshold, no equilibrium binds over exactly the forced window at some K ; these gaps are labeled “no equilibrium” in the figure (see Table 7). In the legends, “restr.” denotes the calibration with attention set to the restricted estimate (§5.7).

4.5 The cost of the lower bound

Does attention change how costly the lower bound is? Consider a demand shock just large enough to hold the economy at the bound for N quarters, and compare the cumulative output loss over ten years with and without the bound (Figure 8; the full set of numbers is in Table OA.8 of the online appendix). The size of the shock is chosen to produce each target

Table 7: Calendar forward guidance at the lower bound: output and inflation today relative to $K = 1$, by extra quarters K at the floor, in an eight-quarter episode. Last column: for a permanent peg, the largest forward eigenvalue when a bounded solution exists, otherwise “indet.” with the average per-quarter growth of the effect. Above the threshold the pegged system has no unique bounded solution (a local test). “n.e.”: no equilibrium keeps the rate at the floor over exactly the forced window (the rate would fall below the floor before the episode or after lift-off). At the calibration, this occurs for $10 \leq K \leq 23$ at $\rho = 0.9$ and for $11 \leq K \leq 23$ at $\rho = 0.897$ (displayed: $K = 12\text{--}20$). At the restricted estimate, it occurs for $19 \leq K \leq 24$ at $\rho = 0.8$ (displayed: $K = 20, 24$). Each displayed entry is the solution with binding set exactly the forced window (one linear system, with a single solution when nonsingular). A broader, non-exhaustive search also allows one extra binding block before the episode and one after the window. At the calibration, it finds candidate binding patterns other than the forced window for every K at $\rho = 0.9$ and $\rho = 0.897$, and for $K \geq 14$ at $\rho = 0.8$. At the restricted estimate, it finds them for every K at $\rho = 0.8$. These patterns bind outside the forced window, and the search implies neither uniqueness nor a count of equilibria. “restr.”: the common attention of the restricted estimate (§5.7), other parameters at the calibration. “Posterior modes”: the rational-expectations mode and the mode of the common-attention restriction; “CD mode, $\rho = \cdot$ ” holds the other parameters at the latter and varies ρ .

extra quarters at the floor K	1	4	8	12	16	20	24	permanent peg: λ_f
<i>Calibration</i>								
output today, relative to $K=1$								
RE	1.00	163.86	4635.29	1.9×10^4	2.9×10^4	7.2×10^4	2.8×10^5	indet. (growth 1.25)
CD, $\rho = 0.9$	1.00	24.13	446.46	n.e.	n.e.	n.e.	4.2×10^4	indet. (n.e. for $10 \leq K \leq 23$)
CD, $\rho = 0.8$	1.00	22.96	143.13	518.97	1566.91	4454.27	1.3×10^4	indet. (growth 1.31)
CD, $\rho = 0.6$	1.00	23.24	52.58	67.40	73.29	75.39	76.11	0.754
CD, $\rho = 0.897$ (restr.)	1.00	23.97	424.76	n.e.	n.e.	n.e.	4.7×10^4	indet. (n.e. for $11 \leq K \leq 23$)
inflation today, relative to $K=1$								
RE	1.00	166.24	4712.83	1.9×10^4	3.0×10^4	7.3×10^4	2.9×10^5	
CD, $\rho = 0.9$	1.00	24.42	453.04	n.e.	n.e.	n.e.	4.3×10^4	
CD, $\rho = 0.8$	1.00	23.27	145.58	528.19	1595.03	4534.44	1.3×10^4	
CD, $\rho = 0.6$	1.00	23.80	54.18	69.55	75.65	77.83	78.58	
CD, $\rho = 0.897$ (restr.)	1.00	24.25	431.02	n.e.	n.e.	n.e.	4.8×10^4	
<i>Posterior modes (CD: common-attention restriction)</i>								
output today, relative to $K=1$								
RE (RE mode)	1.00	58.18	763.03	3735.93	8946.93	1.9×10^4	4.8×10^4	indet. (growth 1.24)
CD, $\hat{\rho} = 0.897$ (CD mode)	1.00	25.05	318.05	2284.41	6646.51	1.3×10^4	3.0×10^4	indet. (growth 1.24)
CD mode, $\rho = 1$ (RE)	1.00	33.12	286.88	1124.48	3519.14	1.1×10^4	3.7×10^4	indet. (growth 1.34)
CD mode, $\rho = 0.90$	1.00	25.20	316.82	2153.20	6131.26	1.3×10^4	3.0×10^4	indet. (growth 1.25)
CD mode, $\rho = 0.80$	1.00	61.91	566.38	3331.61	2.4×10^4	n.e.	n.e.	indet. (n.e. for $19 \leq K \leq 24$)
CD mode, $\rho = 0.60$	1.00	7.09	15.01	19.76	22.09	23.12	23.55	0.768
inflation today, relative to $K=1$								
RE (RE mode)	1.00	58.84	773.00	3784.70	9063.32	1.9×10^4	4.8×10^4	
CD, $\hat{\rho} = 0.897$ (CD mode)	1.00	25.08	318.46	2287.26	6654.75	1.3×10^4	3.0×10^4	
CD mode, $\rho = 1$ (RE)	1.00	33.34	289.01	1132.84	3545.17	1.1×10^4	3.8×10^4	
CD mode, $\rho = 0.90$	1.00	25.23	317.22	2155.89	6138.90	1.3×10^4	3.0×10^4	
CD mode, $\rho = 0.80$	1.00	61.99	567.12	3335.92	2.4×10^4	n.e.	n.e.	
CD mode, $\rho = 0.60$	1.00	7.09	15.01	19.75	22.08	23.11	23.54	

duration; given the shock, the bound binds only when the shadow rate falls below it, so the duration is an equilibrium outcome of each calibrated experiment rather than imposed on agents.

Two results stand out. First, in this model the bound itself costs little under either expectation scheme. Even a twelve-quarter episode raises the output loss by at most about 22% relative to the unconstrained path, and by 3–5% at the posterior modes. This pattern is consistent with habit, price and wage indexation, and the credit-spread block spreading the shock over time, so that the missing policy response is a small part of the total. Second, attention does not systematically flatten the cost. At the calibration, cognitive discounting *raises* the amplification from the bound (1.07 under rational expectations at $N = 12$ against 1.22 at $\rho = 0.897$); at the posterior modes (the rational mode and the restricted estimate) it lowers it (1.050 against 1.026). In absolute terms the extra loss due to the bound is smaller under cognitive discounting at the restricted estimate, and at the calibration for $\rho \leq 0.8$.

This differs from Gabaix’s lower-bound result (his Proposition 3.2 and Figure 1), where in a linear experiment with an exogenous trap length the rational output loss grows without bound as the trap lengthens and sufficiently strong myopia bounds it. Here the trap length is endogenous and, over the durations computed (up to twelve quarters), the structural frictions keep the rational cost small, so the behavioral correction has little to act on; we draw no inference about episodes of unbounded length. The exercise describes the model’s lower-bound mechanism; it is not a data-identified magnitude.

4.6 Fiscal policy: cognitive discounting raises the government-spending multiplier

Alongside forward guidance, the other classic behavioral implication of cognitive discounting is fiscal (Gabaix, 2020; Lubik and Marzo, 2021): in that literature, agents who under-react to the future — to the higher taxes and interest rates a spending expansion eventually implies — crowd out private demand *less*, so the government-spending multiplier is *larger* under bounded rationality. The laboratory model has the lever ($G_t = g_t Y_t$, with a shock to the government-spending share g_t) but no government debt, tax, or financing block, so the future-tax channel is the motivation from the literature, not a mechanism of our model. What the model contains is the forward-looking channel through the household’s discounted expectations of future marginal utility and of the real interest rate that policy sets in response to the expansion. Table 8 and Figure 9 report the output multiplier dY/dG — at impact and in present value — in normal times and inside an eight-quarter lower-bound episode.

The pattern is the one the behavioral literature predicts, and it is sizable. At the cali-

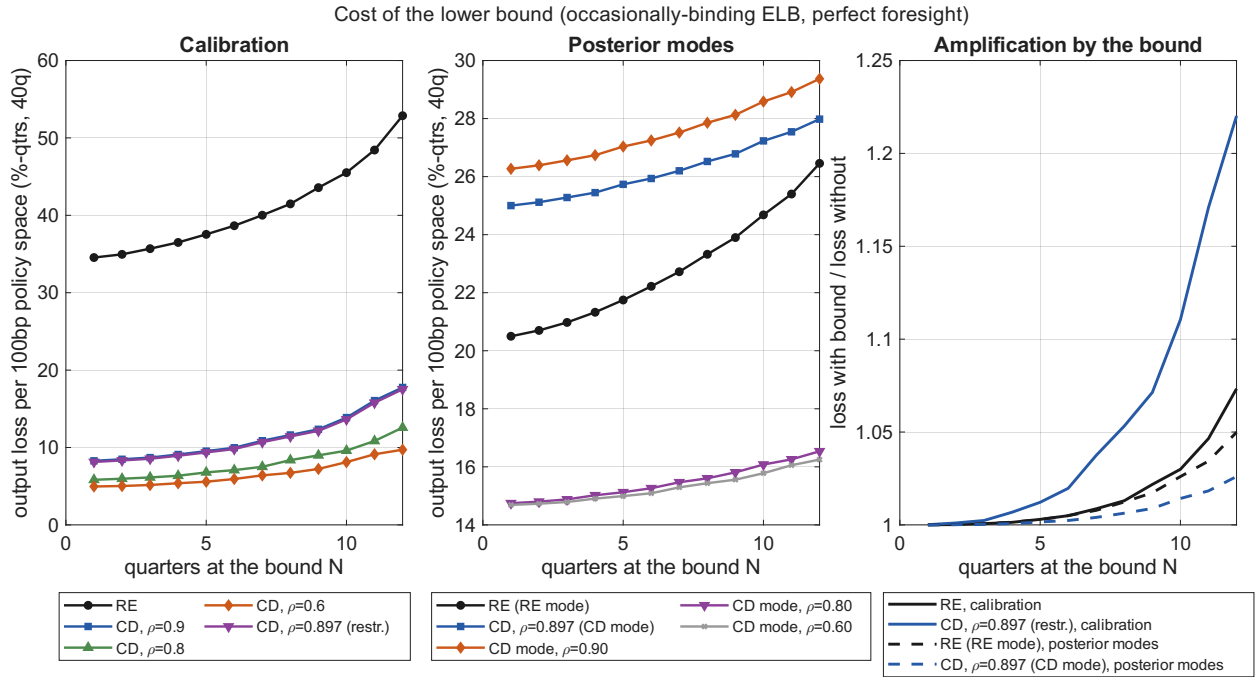


Figure 8: Cost of the lower bound (perfect foresight). Left and center: cumulative output loss over 40 quarters per 100bp of policy space, against the number of quarters N the bound binds (calibration; posterior modes, with the restricted estimate of §5.7 for cognitive discounting). Right: loss with the bound relative to the loss without it. In the legends, “restr.” denotes the calibration with attention set to the restricted estimate.

bration, in normal times, the impact multiplier rises from 0.80 under rational expectations to 0.86 with attention set to the restricted estimate, $\rho = 0.897$, and to 0.96 at $\rho = 0.6$. The present-value multiplier, which cumulates the weaker crowding-out, rises more, from 0.45 to 0.57 and 0.80. At the bound, where the policy rate does not respond, the multiplier is larger under both schemes and attention again raises it: the impact multiplier goes from 0.91 to 1.00 at $\rho = 0.897$. The same ordering holds at the posterior modes (present value 0.34 at the rational mode, 0.48 at the restricted mode; at the bound 0.40 against 0.68). The robust content is the direction: discounting the future makes government spending more potent, as in the behavioral fiscal result of [Gabaix \(2020\)](#), here obtained through the expectational channels the nonlinear, nonstationary, estimated model contains rather than through a failure of Ricardian equivalence. The experiments differ in detail: his Proposition 5.1 shows that budget deficits enter the behavioral IS curve, a failure of Ricardian equivalence, and a deficit-financed purchase at a fixed nominal rate then has an impact multiplier above one. Ours is a spending shock under an active Taylor rule with habit and credit frictions, consistent with our levels sitting near or below one.

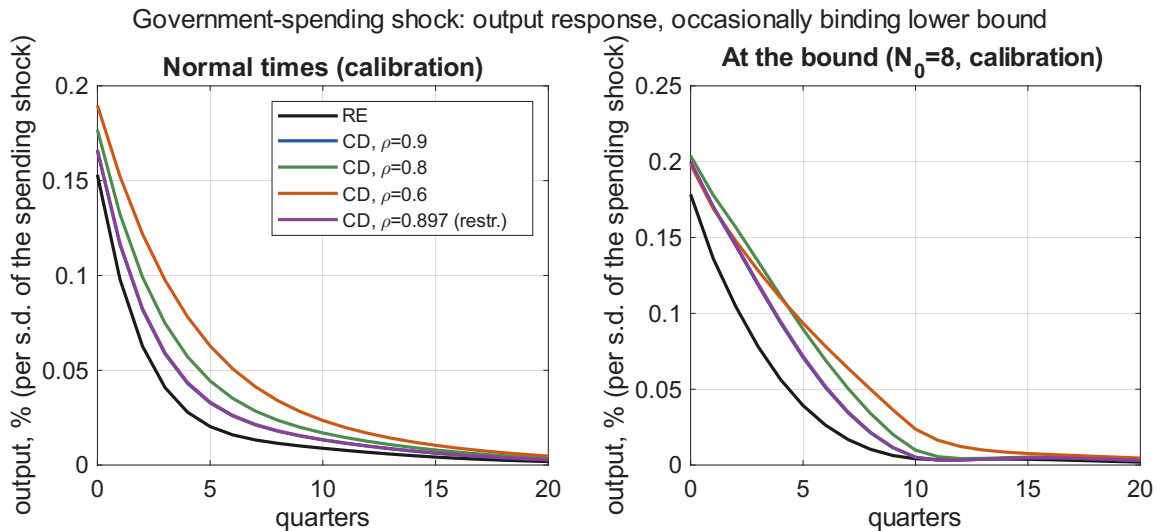


Figure 9: Government-spending multiplier. Output response to a government-spending-share shock in normal times (left) and inside an eight-quarter lower-bound episode (right), at the calibration, under rational expectations and cognitive discounting. In the legend, “restr.” refers to the common attention of the restricted estimate (§5.7).

4.7 Optimal policy in the nonlinear model, order by order

The operator also lets us pose *optimal* monetary policy under bounded rationality directly in the fully nonlinear model at successive perturbation orders (here through third). We use

Table 8: Output multiplier dY/dG of a government-spending shock, by attention ρ : impact and present value (discount factor 0.99, 40 quarters), in normal times and inside an eight-quarter lower-bound episode, for a marginal and a one-standard-deviation shock. The first column is the forward-propagation eigenvalue away from the bound. “restr.”: the common attention of the restricted estimate, other parameters at the calibration; “Posterior modes”: the rational-expectations mode and the mode of the common-attention restriction (§5.7).

		normal times		at the bound (marginal)		at the bound (1 s.d.)	
	$\max \lambda(\mathcal{F}) $	impact	PV	impact	PV	impact	PV
<i>Calibration</i>							
RE	0.925	0.803	0.448	0.912	0.593	0.912	0.593
CD, $\rho = 0.9$	0.869	0.859	0.566	0.999	0.788	0.999	0.788
CD, $\rho = 0.8$	0.809	0.905	0.665	1.015	0.884	1.015	0.884
CD, $\rho = 0.6$	0.647	0.959	0.804	0.992	0.933	0.987	0.926
CD, $\rho = 0.897$ (restr.)	0.867	0.860	0.569	1.000	0.792	1.000	0.792
<i>Posterior modes (CD: common-attention restriction)</i>							
RE (RE mode)	0.913	0.823	0.341	0.878	0.395	0.878	0.395
CD, $\hat{\rho} = 0.897$ (CD mode)	0.917	0.828	0.477	0.998	0.683	0.993	0.676
CD mode, $\rho = 1$ (RE)	0.963	0.711	0.335	0.703	0.326	0.703	0.326
CD mode, $\rho = 0.90$	0.918	0.825	0.472	0.993	0.675	0.990	0.670
CD mode, $\rho = 0.80$	0.858	0.902	0.604	1.034	0.840	1.003	0.802
CD mode, $\rho = 0.60$	0.679	0.966	0.783	0.969	0.879	0.969	0.869

a *stationary policy variant* of the laboratory model. It keeps the private block (internal habit, a credit-spread Euler wedge, sticky prices *and* wages with reference indexation, and the cognitive-discounting operator applied to the exact nonlinear conditions). It changes three things. Technology follows a stationary autoregressive process instead of a unit root with drift, so the variant is a different model rather than a detrending of the laboratory model. The price level is replaced by inflation, so that the planner’s Lagrange multipliers are constant. And there is no lower bound: the net policy rate is the planner’s instrument. The operator discounts the same forward-looking variables except the price level, which no longer appears. The planner’s objective is the representative household’s period utility $\zeta_{C,t} \log(C_t - bC_{t-1}) - \zeta_{H,t} H_t^{1+\nu} / (1 + \nu)$, the same internal-habit utility that generates the private conditions, discounted at the *true* discount factor. The planner’s conditions are treated exactly like the private agents’ conditions. We first derive the first-order conditions of the Ramsey (commitment) problem under rational expectations. The operator is then applied to every expectation in the resulting system, the planner’s conditions included (Online Appendix OA-G.13). The planner therefore forms expectations as the private agents do. In Gabaix (2020) and Benchimol and Bounader (2023), by contrast, the central bank is rational and only private agents discount the future. Because optimal policy is a property of the economy, not of a sample, we evaluate it at the *baseline calibration* used for the other model-property results above — deliberately *decoupled from estimation*.

The commitment problem is determinate at all 22 values of ρ we check on $[0.1, 1]$, at this calibration and at the restricted estimate. The Blanchard–Kahn root count is exactly balanced, with 19 stable roots for 19 predetermined variables and no unit root. The problem solves at first, second and third order for rational expectations and for cognitive discounting. Cognitive discounting reshapes optimal policy already at first order: the optimal responses to a price-markup or credit-spread shock are attenuated relative to rational expectations. The peak responses fall by about a quarter for the output gap and by roughly 5 to 7 percent for inflation (computed from peaks reported to three significant digits; Figure 10). At this calibration the second-order corrections to optimal inflation are small in this experiment: the pruned second-order inflation response differs from the first-order one by about 10^{-3} of its peak, and the risky steady state of inflation moves by less than 10^{-4} percentage points at a quarterly rate (Table 9). The first measures the deterministic second-order correction to the impulse response; the second measures the uncertainty-induced shift in the steady state. Both are small in this experiment, which does not establish that risk cannot matter for optimal policy or welfare in other experiments or models.

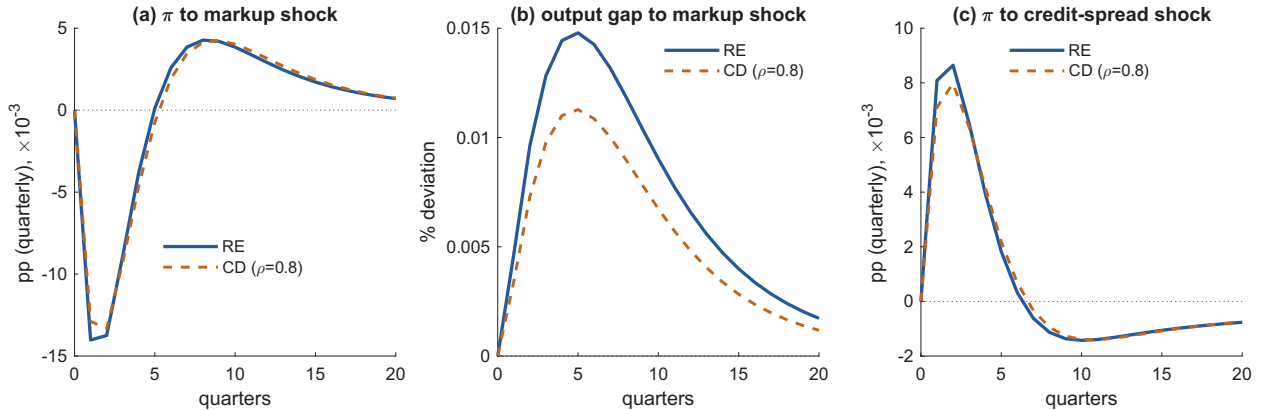


Figure 10: Optimal (Ramsey, commitment) policy at the baseline calibration, welfare objective, first order, with the operator applied to the whole optimality system. Impulse responses under rational expectations (solid) and cognitive discounting $\rho = 0.8$ (dashed): (a) inflation and (b) the output gap to a price-markup shock; (c) inflation to a credit-spread shock. Inflation in percentage points (quarterly rate); output gap in percent. The problem solves at orders 1–3 under both expectation schemes and is determinate at both attention levels.

Discretion. We solve the discretionary (Markov-perfect) problem on the same stationary model. At the same baseline calibration, and at both first and second order, discretion solves cleanly for rational expectations and for cognitive discounting. Welfare is measured as $W \equiv \mathbb{E}[\zeta_{C,t} \log(C_t - bC_{t-1}) - \zeta_{H,t} H_t^{1+\nu}/(1+\nu)]$, the unconditional mean of the period utility flow (not lifetime utility), computed from the second-order solution by sim-

Table 9: Second-order corrections to optimal commitment: peak of the deterministic second-order correction to the inflation impulse response (pruned order-2 minus order-1 response, variance terms excluded) relative to the peak order-1 response, and the uncertainty-induced shift of the steady state of inflation (risky minus deterministic steady state).

	peak order-1 response of π (pp, quarterly)	peak order-2 correction / peak response	risky-steady-state shift of π (pp, quarterly)
RE	0.0140	1.1×10^{-3}	7.5×10^{-6}
CD ($\rho = 0.8$)	0.0133	1.2×10^{-3}	1.4×10^{-5}

ulation with common random numbers across all cases. The cost of discretion is reported as $100(W^{\text{commitment}} - W^{\text{discretion}})$. With log utility in the habit-adjusted consumption surplus, a permanent proportional cut c in consumption (current and lagged alike) lowers W by $-\bar{\zeta}_C \log(1 - c)$, where $\bar{\zeta}_C \equiv \mathbb{E}[\zeta_{C,t}]$ is the unconditional mean of the stochastic preference weight (normalizing its steady-state value to one does not set this mean to one). The exact per-period consumption equivalent is therefore $1 - \exp\{-(W^{\text{commitment}} - W^{\text{discretion}})/\bar{\zeta}_C\}$. The reported number is the approximation $c \approx W^{\text{commitment}} - W^{\text{discretion}}$, which linearizes the exponential and sets $\bar{\zeta}_C \approx 1$. Linearization accounts for less than 0.0005 percentage points at these magnitudes, so the exact and the reported values can differ by one unit in the third reported decimal (0.307 against 0.306 under rational expectations). The normalizer, $\bar{\zeta}_C = 1 + 4.5 \times 10^{-5}$ for the calibrated log-AR(1) weight, moves the equivalent by about 0.00001 percentage points. Two results follow (Figure 11, Table 10). First, commitment dominates discretion at every attention level we compute: the unconditional welfare cost of discretion is 0.31% of consumption under rational expectations. Second, cognitive discounting *erodes the value of commitment*, and does so monotonically in attention on the computed grid — the cost of discretion falls to 0.292% at $\rho = 0.8$ and 0.283% at $\rho = 0.6$. Because the operator also discounts the planner’s own conditions, this erosion reflects the planner’s and the private agents’ discounting together. The exercise does not isolate how private agents respond to promises. The erosion is small: at $\rho = 0.9$, close to the estimated common attention, the cost is 0.298%, and even at $\rho = 0.1$ it is still 0.269%. Discretion carries a large average-inflation bias (mean inflation of about 9.7% a year with rational expectations), which attention changes only slightly; we do not decompose the cost of discretion into this bias and stabilization. The linear antecedent is again [Gabaix \(2020\)](#) (his Proposition 4.2): with behavioral agents, optimal commitment no longer implies price-level targeting, because promises about the far future lose leverage over myopic agents. [Benchimol and Bounader \(2023\)](#) extend the analysis across a menu of myopias: inflation targeting is optimal when myopia distorts inflation expectations, price-level targeting under the other myopias they consider. Our result points in the same direction — the value of commitment itself erodes

— now at second order, with the risk channel active, and with a planner that shares the agents' expectations. The erosion is modest, as in the other exercises, consistent with damping by the model's habit and credit frictions. At first order, all four tested initial guesses converged to the same discretionary decision rule; no alternative solution was found in that probe, which is evidence about the search, not a uniqueness result.

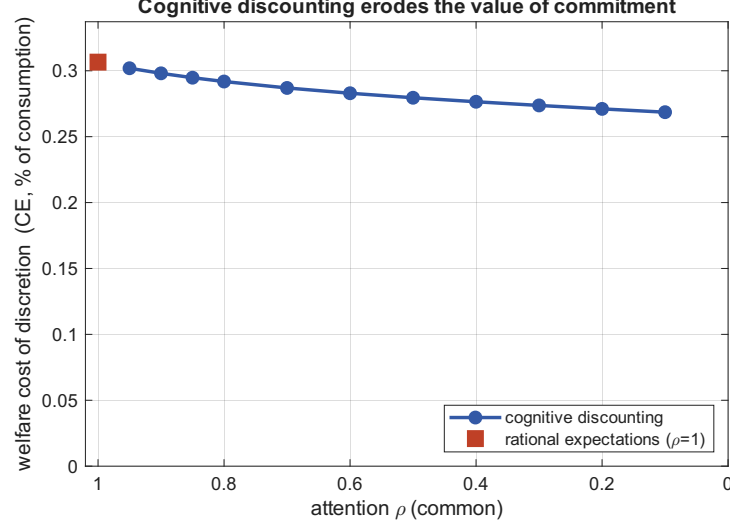


Figure 11: The welfare cost of discretion (consumption equivalent, percent) by attention ρ : 100 ($W^{\text{commitment}} - W^{\text{discretion}}$), with W the unconditional mean period utility at second order on the stationary policy variant, common random numbers across all cases. Commitment dominates at every ρ ; cognitive discounting erodes its value monotonically on the grid.

Table 10: Commitment versus discretion by attention ρ : unconditional expected welfare (mean period utility W , order 2) and the consumption-equivalent cost of discretion, 100 ($W^{\text{commitment}} - W^{\text{discretion}}$), on the stationary policy variant.

ρ	$\mathbb{E}[W]$ commitment	$\mathbb{E}[W]$ discretion	cost of discretion (CE %)	share of RE cost	$\mathbb{E}[\pi]$ discretion (ann. %)
1.00 (RE)	-2.832414	-2.835479	0.307	1.00	9.68
0.95	-2.832415	-2.835434	0.302	0.98	9.62
0.90	-2.832420	-2.835401	0.298	0.97	9.57
0.85	-2.832428	-2.835375	0.295	0.96	9.53
0.80	-2.832436	-2.835355	0.292	0.95	9.49
0.70	-2.832454	-2.835324	0.287	0.94	9.43
0.60	-2.832471	-2.835301	0.283	0.92	9.38
0.50	-2.832486	-2.835281	0.280	0.91	9.34
0.40	-2.832499	-2.835264	0.276	0.90	9.31
0.30	-2.832510	-2.835247	0.274	0.89	9.27
0.20	-2.832520	-2.835231	0.271	0.88	9.24
0.10	-2.832529	-2.835215	0.269	0.88	9.21

5 Estimation: methodology and results

This section illustrates the method in one model with one set of observables. No single estimate in it is a verdict on the method: how attention should differ across variables is specific to the model and the data. The operator is designed for richer models, in which restrictions such as a common or block-specific discount become testable rather than assumed. The baseline estimates attention variable by variable, as the operator is built to do (§§5.4–5.6); restrictions of it are reported for comparison (§5.5) and as a robustness check (§5.7).

5.1 Data, sample, and observables

The model is taken to US data over **1985Q1–2019Q4** (pre-COVID), with seven observables: the net policy rate (level), the inflation gap, the credit spread (level), and the demeaned real growth rates of output, consumption, the real wage, and government spending. Each observable carries an independent measurement error. The series are from FRED. The interest-rate and price series are the effective federal funds rate (FEDFUNDS), the GDP deflator (GDPDEF), and Moody’s Baa yield relative to the 10-year Treasury (BAA10YM); inflation enters as quarterly inflation minus a 0.5% quarterly target. The quantities are real GDP (GDPC1), real personal consumption expenditures (PCECC96), real compensation per hour in the nonfarm business sector (COMPRNFB), and real government consumption expenditures and gross investment (GCEC1). They enter as demeaned quarterly log growth rates. The policy-shock standard deviation is fixed at its calibrated value. All computations are implemented in RISE.

5.2 Priors

Each of the eight discounted variables, $X \in \{C, \Lambda, P, Y, \pi, \pi^w, \tilde{\pi}, \tilde{\pi}^w\}$, has its own attention ρ_X , and each attention receives a flat Uniform(0, 1) prior, independent across variables. The flat prior lets the data place each attention anywhere in the unit interval, and it gives the sector-specific and block-specific restrictions of §5.5 the same prior family, so that their marginal likelihoods are comparable. Structural parameters use the usual Beta (persistences, indexation, habit), Gamma (inverse Frisch elasticity, Taylor-rule coefficients), and inverse-gamma (shock standard deviations) families. The standard deviations of the measurement errors are estimated, with inverse-gamma priors whose means are about 40% of each series’ sample standard deviation. The robustness restriction of §5.7 uses a Beta prior on its single attention parameter, with four alternatives (§5.7.2). Appendix A lists the priors.

5.3 Estimation strategy

The model is estimated in its nonstationary form, with no detrending by the econometrician. Three features make this possible. First, the operator’s reference point for a growing variable moves with its trend (Section 2), so cognitive discounting applies to output, consumption, and the price level exactly as to stationary variables. Under rational expectations the levels model and its detrended twin give the same likelihood to machine precision (Online Appendix OA-G.9). Second, the diffuse component of the state, the unit roots, is initialized with an exact diffuse prior (Koopman and Durbin, 2003; Durbin and Koopman, 2012), so no presample is discarded and no arbitrary initial variance enters. Exactness refers to this linear initialization, not to the filter as a whole. Third, the lower bound enters the likelihood as the occasionally binding constraint of Section 3, filtered with a constraint-aware filter in the spirit of Binning and Maih (2015), which is an approximation; §5.7.5 compares the estimates without the bound. The reported likelihoods, and the evidence ranking built on them, are conditional on that approximate filter.

Two further choices matter, and both are documented in Online Appendix OA-E. All measurement errors are estimated rather than fixed at a negligible value, since with near-zero measurement errors a handful of crisis quarters can receive disproportionate weight in the likelihood. And because the likelihood of the behavioral model has several local peaks, the posterior mode is located with a global optimizer (an artificial bee colony; Karaboga and Basturk, 2007), warm-started along the chain of nested models. Posteriors are sampled by Metropolis–Hastings in logit-transformed parameters. For the attention specifications of Table 11, the sampler uses a fixed mixture of local random-walk, independence and mode-jump proposals, with four chains started from dispersed points. Marginal data densities are computed from the posterior samples (bridge sampling, Meng and Wong, 1996, and the modified harmonic mean, Geweke, 1999) and, where the mode is interior, by the Laplace approximation with a finite-difference Hessian. Convergence is assessed by the rank-normalized split- $\hat{R}$ and the bulk and tail effective sample sizes of Vehtari et al. (2021); the standard we apply is $\hat{R} < 1.01$ and both effective sample sizes at least 400.

5.4 Attention estimated variable by variable

Table 11 reports the baseline: every discounted variable has its own attention, on the nonstationary model, data and lower-bound filter just described. Every chain in the table meets the convergence standard: across all parameters and the six specifications of panel B, rational expectations included, the largest $\hat{R}$ is 1.009 and the smallest bulk or tail effective sample size 428 (442 over the five behavioral specifications).

Posterior mean attention is uneven across variables (Table 11, panel A). The posterior means are 0.89 for consumption, 0.91 for price inflation, 0.76 for marginal utility, and 0.79 for wage inflation. Their posterior standard deviations are between 0.08 and 0.14, against 0.29 under the uniform prior, so the data are informative about these four attentions. The posterior of marginal-utility attention lies below one (95% HPD interval [0.62, 0.90]); those of consumption, price-inflation and wage-inflation attention lie mostly in the upper half of the unit interval, and their 95% HPD intervals reach one to the displayed precision. A posterior mean below one is automatic under a continuous prior on the unit interval, so it does not by itself measure a departure from rational expectations; that question is answered by comparing marginal likelihoods (§5.5). The two reference inflation rates that index prices and wages have lower posterior means, 0.52 (prices) and 0.45 (wages), and wide posteriors (standard deviations 0.28 and 0.26); freeing the two jointly is positively favored over fixing both at one (§5.5). Under the indexation rules of Appendix A, next period’s reference inflation rates are set by this period’s inflation, so they are known when expectations are formed. Attention to them therefore measures how agents weight an indexation benchmark that is already determined, not their response to uncertainty about it. Attention to the price level and to output is not identified at first order and integrates out (§5.6). The posterior modes lie on a bound of the unit interval for some attentions, where the Hessian is not positive definite, so the Laplace approximation is not available and the marginal likelihoods come from the posterior samples. Under a Beta(0.85, 0.10) prior on every attention instead, the posterior mode has all eight attentions interior, between 0.85 and 0.97 (mode only; Online Appendix Table OA.3).

5.5 How the data rank the expectation schemes

The literature attaches attention to agents — households and firms in Gabaix (2020) and Hirose et al. (2024) — or, more finely, to individual forward-looking conditions. We estimate two such alternatives on the same model, data and flat priors. The sector-specific version has one attention for the household variables (consumption and marginal utility) and one for the remaining discounted variables (prices, inflation, wages, and output), each applied wherever those variables appear. It groups variables rather than giving agents distinct beliefs about a shared variable, so it is a restriction of the operator. The block-specific version has one attention for each forward-looking condition (the consumption first-order condition, the Euler equation, and the wage and price Phillips curves). A variable that appears in several conditions then receives a different attention in each, so this version is not in general a restriction of the variable-specific operator. At first order in this model it is one. The

Table 11: Attention estimated variable by variable, and its restrictions, with flat priors. Panel A: posterior of the baseline attentions. Panel B: sampled log marginal data density, with the batch-dispersion diagnostic of §5.5 in parentheses, and convergence diagnostics over all parameters; every row meets $\hat{R} < 1.01$ and effective sample sizes ≥ 400 . Posterior modes and a Beta-prior variant: Online Appendix Table OA.3.

<i>A. Baseline: posterior of each attention, Uniform(0,1) priors</i>					
Attention on		mean	s.d.	95% HPD interval	
Consumption	ρ_C	0.89	0.13	[0.62, 1.00]	
Marginal utility	ρ_Λ	0.76	0.08	[0.62, 0.90]	
Price inflation	ρ_π	0.91	0.09	[0.74, 1.00]	
Wage inflation	ρ_{π^w}	0.79	0.14	[0.53, 1.00]	
Reference price inflation	$\rho_{\bar{\pi}}$	0.52	0.28	[0.06, 1.00]	
Reference wage inflation	$\rho_{\bar{\pi}^w}$	0.45	0.26	[0.00, 0.88]	
Price level, output	ρ_P, ρ_Y	not identified at first order; integrated out			

<i>B. Sampled log marginal data density by attention structure, flat priors</i>					
Specification	sampled attentions	bridge	MHM	max $\hat{R}$	min ESS bulk/tail
Variable by variable (baseline)	6	4189.93 (0.10)	4190.80 (0.13)	1.006	751/442
with $\rho_{\bar{\pi}} = \rho_{\bar{\pi}^w} = 1$	4	4188.53 (0.12)	4188.87 (0.30)	1.009	878/668
Households/firms	2	4190.29 (0.18)	4190.75 (0.21)	1.009	881/950
Block-specific	4	4188.88 (0.18)	4189.17 (0.45)	1.009	831/461
Common ρ (§5.7) [†]	1	4188.50 (0.09)	4188.54 (0.10)	1.007	894/1242
Rational expectations [†]	0	4175.7 (0.2)	4175.9 (0.3)	1.004	1149/428

[†]Sampled with an earlier development version of the software than the other rows (§5.5). Rational expectations: the sample of Table 13 (same priors on all common parameters; random-walk Metropolis). Panel A: HPD bounds are rounded; 1.00 is not an atom. The price-level and output attentions are held fixed in sampling, which leaves the marginal likelihood unchanged (§5.6).

variables shared across conditions (marginal utility, the price level and output) enter the Phillips curves' forward terms only multiplied by indexation gaps that vanish in the steady state. The first-order block-specific model is therefore the variable-specific one with ρ_C equal to the attention of the consumption condition, ρ_Λ to that of the Euler equation, $\rho_\pi = \rho_{\bar{\pi}}$ to that of the price Phillips curve and $\rho_{\pi^w} = \rho_{\bar{\pi}^w}$ to that of the wage Phillips curve. The estimation is at first order; beyond it the block-specific version is an alternative specification rather than a restriction. Both use the same moving reference as the operator. A third restriction keeps attention variable by variable but imposes rational expectations on the two reference inflation rates, $\rho_{\bar{\pi}} = \rho_{\bar{\pi}^w} = 1$. Evidence is graded on the scale of Kass and Raftery (1995), in terms of twice the log Bayes factor, $2 \ln B$: below 2 it is not worth more than a bare mention, between 2 and 6 it is positive, between 6 and 10 strong, and above 10 very strong.

Cognitive discounting beats rational expectations under every attention structure, and the evidence is very strong. The bridge-sampling log marginal data density of the variable-by-variable baseline is 4189.9, against 4175.7 for rational expectations (Table 11; for rational expectations, which has no attention parameter and shares every other prior, Table 13), a margin of 14.2 log points ($2 \ln B \approx 28$). The variable-specific, households/firms and block-specific specifications, with and without the reference-inflation restriction, were sampled with one fixed version of the software; the rational-expectations comparator and the common-attention restriction were sampled earlier, with development versions whose exact revision was not recorded. Re-evaluated with the fixed version, the log posteriors at two stored modes are reproduced exactly: the rational-expectations mode and the common-attention mode under the Beta prior of §5.7. This checks those two points, not the whole posterior or the marginal likelihood. Across the five attention structures of Table 11 the margin ranges from 12.8 to 14.6 log points: a Bayes factor above e^{12} in every case, far beyond the threshold for very strong evidence.

Freeing the two reference-inflation attentions is positively favored. These attentions have the lowest posterior means, 0.45 and 0.52, and wide posteriors (§5.4). Fixing both at one lowers the log marginal data density to 4188.5, 1.40 log points below the baseline: a Bayes factor of about 4.1 ($2 \ln B \approx 2.8$), positive evidence in favor of freeing the two attentions jointly. Under equal prior odds it gives the unrestricted model a posterior probability of about 80 percent. The comparison concerns fixing both attentions jointly; it does not test either of them separately against one. The common, sector-specific and block-specific restrictions estimated here tie the reference inflation rates to other variables, or to the other forward terms of the Phillips curves in which they appear, so they cannot reveal this distinction; a grouping with its own reference-inflation attention could. Estimating attention variable by

variable reveals it without the grouping having to be chosen in advance.

Which restrictions the data accept. Attention by variable is favored over a common attention by 1.43 log points (a Bayes factor of about 4.2, $2 \ln B \approx 2.9$; §5.7), which is positive evidence. It is favored over the block-specific restriction by 1.05 log points (about 2.9, $2 \ln B \approx 2.1$), which is positive but borderline. The two-group households/firms restriction fits about as well in this model. Bridge sampling puts it 0.36 log points above the baseline and the modified harmonic mean 0.05 below, a Bayes factor of at most about 1.4 either way ($2 \ln B \approx 0.7$), a difference in evidence not worth more than a bare mention. The modified harmonic mean gives the same evidence grades throughout, with larger margins over the block-specific, common and restricted-reference-inflation versions. That the coarser households/firms grouping costs nothing here is learned by estimating the variable-by-variable model and comparing: the framework lets such restrictions be evaluated rather than imposed in advance. In another model, or with other observables, the right grouping may differ, which is why the method is built variable by variable.

Numerical precision. The figures in parentheses in Table 11 are batch-dispersion diagnostics, not calibrated errors of the pooled estimates. For the five behavioral specifications, each of the four chains is split into two halves, each half with its own fitted auxiliary density, and the diagnostic is the standard deviation of the eight half-chain estimates divided by $\sqrt{8}$. The rational-expectations row uses six half-chain batches from its three chains. In each of the five behavioral rows, the pooled bridge estimate exceeds the mean of its half-chain estimates by 0.47 to 0.76 log points, more than the batch dispersion. This is consistent with an effect of sample size on the estimator, which the stored results do not separate from the effect of refitting the auxiliary density. Among the behavioral specifications, the differences between the baseline and each restriction computed from the half-chain means have the same signs as those from the pooled estimates and lie within 0.29 log points of them. These shifts are small relative to the reported margin of about 14 log points over rational expectations, but this batch comparison does not quantify the error in that margin. Differences of a few tenths of a log point among attention structures, such as that between the baseline and the households/firms grouping, are not numerically resolved by these diagnostics.

Scope. These estimates illustrate the method in one model with one set of observables. The operator is designed for richer models, in which restrictions such as these become testable rather than assumed.

The restrictions are informative about *where* the inattention sits within their groupings. At their posterior modes under Beta(0.85, 0.10) priors, the estimated inattention is in the household Euler equation ($\rho_{\text{Euler}} = 0.854$; households $\rho_H = 0.868$), while firms and wage setters are close to rational (0.95–0.97, within about one standard error of one; Online

Appendix Table [OA.7](#)). This matches the finding of rational price setters in [Hohberger et al. \(2024\)](#), in a linearized estimated model. The firm-side attentions are nearly orthogonal to everything else and only weakly identified. These are results of the restrictions, under Beta priors. In the variable-by-variable baseline, attention to wage inflation (0.79 in posterior mean) has the second-lowest posterior mean of the four attentions the data inform most. Attaching one attention to a forward-looking condition is not the same as attaching it to a variable, so the near-rational wage setters of the block-specific restriction do not carry over to the baseline.

5.6 What identifies attention: the model and the observables

Which attentions are identified depends on the model and on the observables. The marginal likelihood compares prior-averaged fit; it imposes no automatic penalty on parameters the data do not pin down, and parameters that are completely likelihood-inactive integrate out without penalty under independent normalized priors. With the present observables, the posteriors of the attentions on consumption, price inflation, marginal utility and wage inflation are much tighter than the prior. The two reference-inflation attentions have wide posteriors, though fixing both at one is positively disfavored. The price-level and output attentions are not identified at first order. Other observables, such as survey expectations, could identify more of them.

The last property is exact. The operator contracts the deviation of a variable dated $t + 1$ from its reference point, so it moves the first-order solution only through the derivative of the forward-looking conditions with respect to that variable. In this model P_{t+1} and Y_{t+1} appear only in the forward terms of the wage and price Phillips curves, [\(A.7\)](#) and [\(A.13\)](#), where they multiply an indexation gap that is zero in the steady state. Those derivatives are therefore zero, and moving ρ_P or ρ_Y over the unit interval leaves the likelihood unchanged at first order. Their priors are proper, normalized and independent of the remaining parameters θ , and they enter the likelihood through no other channel. They therefore integrate out with no evidence penalty: with $\eta = (\rho_P, \rho_Y)$, $L(y \mid \theta, \eta) = L(y \mid \theta)$ and $\pi(\theta, \eta) = \pi(\theta)\pi(\eta)$ give $\pi(\eta \mid y) = \pi(\eta)$ and $Z = \int L(y \mid \theta)\pi(\theta) d\theta$. They are therefore held fixed in sampling, which leaves the marginal likelihood unchanged; attention to them could be identified, if at all, only at second or higher order ([Online Appendix OA-E.1](#) states the general condition). The other confound usually feared in linearized estimations, between attention and the volatility of the shocks that move the same expectations ([Remark 8](#)), is examined in [§5.7.3](#).

5.7 Robustness: a common-attention restriction

A common attention ρ for all eight variables is a restriction of the operator, reported here as a robustness check only. Its drawback is that it blurs the distinction between the variable-by-variable and the block-by-block strategies. With equal attentions, the variable-specific, sector-specific and block-specific operators coincide, and so do their likelihoods, to machine precision. A common-attention fit therefore cannot tell whether attention attaches to variables or to equilibrium conditions. Under the flat prior of the baseline its marginal likelihood is 4188.5 (bridge), 1.43 log points below the variable-by-variable one (Table 11). With one attention parameter it is cheap to sample under the flat and the Beta(0.85, 0.10) priors, to check without the lower bound, and to use for comparing the operator with term-wise discounting. The checks below use a Beta prior on ρ with mean 0.85 and standard deviation 0.10, centered at the value Gabaix (2020) uses for his calibration; Online Appendix OA-E.2 gives the details.

5.7.1 Estimate and marginal likelihood

Table 12 reports the posterior modes. The common attention is estimated at $\rho = 0.897$ at the mode (standard error 0.023) and at 0.879 in the mean (standard deviation 0.029, 95% HPD interval [0.820, 0.932]). The log marginal data density exceeds that of rational expectations by 14.4 log points (Laplace) and by between 14.4 and 15.3 across the six estimators of Table 13. Besides a Laplace approximation at the mode, five are computed from the posterior samples: the modified harmonic mean (Geweke, 1999), bridge sampling (Meng and Wong, 1996), the Chib and Jeliazkov (2001) estimator, a Laplace approximation around the posterior sample, and an unpublished estimator of Ulrich Müller described in Liu et al. (2011). Their batch-dispersion diagnostics (chains times two halves, as in Table 11) are about 0.5 at most; the spread across estimators (0.9 log points) is one measure of sensitivity, not a calibrated Monte Carlo error bar or a bound on common sampling or approximation error. We run three chains of 220,000 draws for rational expectations and four of 165,994 for cognitive discounting, started in both local modes. All meet the convergence standard: every parameter has $\hat{R}$ below 1.01 (at most 1.004 and 1.006) and bulk and tail effective sample sizes of at least 400. Reciprocal importance sampling (Gelfand and Dey, 1994) and the Sims et al. (2008) estimator also favor cognitive discounting (15.3 and 17.3); the latter agrees on the ranking but not on the magnitude. At the modes, the Laplace margin of 14.4 splits (components rounded separately, from unrounded values) into fit (+9.9 in log-likelihood), prior density (+8.2, most of it on parameters other than ρ), and posterior curvature (−3.6): the priors reinforce rather than create the ranking. A continuous posterior cannot by itself

test the point $\rho = 1$; that test is the comparison of marginal data densities. The posterior has a second, minor mode (about 4% of the mass) that differs from the main one in shock persistences, not in attention (its mean ρ is 0.874).

Table 12: Common-attention restriction: posterior modes under rational expectations (RE) and cognitive discounting with a common attention ρ (CD), with the occasionally binding lower bound and without it. Nonstationary model, first order, 1985Q1–2019Q4. Log marginal data density by the Laplace approximation with a finite-difference Hessian; in brackets, the same with a different step size.

	Lower bound (ELB)		No lower bound	
	RE	CD	RE	CD
$\log L$	4234.2	4244.0	4234.0	4244.6
$\log$ posterior	4285.9	4304.0	4286.1	4304.1
$\log$ MDD (Laplace)	4172.8	4187.2	4173.5	4187.9
second step size	[4172.9]	[4187.3]	[4173.5]	[4188.0]
ρ (mode)	—	0.897	—	0.896
s.e.	—	(0.023)	—	(0.022)

Table 13: Common-attention restriction: log marginal data density across estimators. Random-walk Metropolis in logit-transformed parameters; RE 3 chains $\times$ 220,000 draws, CD 4 chains $\times$ 165,994 draws; every parameter has rank-normalized split- $\hat{R} < 1.01$ and bulk and tail effective sample sizes ≥ 400 . In parentheses, the batch-dispersion diagnostic (batches: chains $\times$ two halves), not a calibrated error of the pooled estimate.

$\log$ MDD	RE	CD	CD – RE
Laplace, finite-difference Hessian at the mode	4172.8	4187.2	14.4
Laplace, posterior-sample covariance	4175.9 (0.4)	4190.5 (0.3)	14.7 (0.5)
Modified harmonic mean	4175.9 (0.3)	4191.0 (0.2)	15.1 (0.3)
Bridge sampling (Meng–Wong)	4175.7 (0.2)	4190.6 (0.1)	14.9 (0.2)
Chib–Jeliazkov	4172.2 (0.1)	4187.5 (0.5)	15.3 (0.5)
Müller	4177.3 (0.3)	4192.0 (0.2)	14.7 (0.4)

5.7.2 Sensitivity to the prior

The Beta prior is informative and centered below one (its 90% interval is roughly [0.66, 0.97] and its density vanishes at $\rho = 1$). Re-estimating under a uniform prior on $[0, 1]$, a prior centered at 0.70, and two priors that push toward rationality, Beta(0.95, 0.03) and Beta(0.99, 0.01), barely moves the estimate: the mode is 0.895 under the uniform prior (against 0.897 under Beta(0.85, 0.10)), 0.890 under the prior centered at 0.70, and 0.913

under $\text{Beta}(0.95, 0.03)$. Only $\text{Beta}(0.99, 0.01)$, which all but imposes rationality, pulls it further, to 0.934. The fit at the mode is the same under every prior (log-likelihoods within ± 0.8); what the prior changes is the marginal data density, and the Laplace approximation favors cognitive discounting over rational expectations under each prior, by 8.7 to 14.4 log points. For the alternative priors we report posterior modes and Laplace approximations; under $\text{Beta}(0.99, 0.01)$ the best point found lies on a parameter bound, so the table reports an interior evaluation point (the best interior point found, not necessarily a stationary point) and its Laplace value (Online Appendix Table [OA.5](#) and Figure [OA.3](#)).

5.7.3 Attention and shock volatility

In linearized estimations, attention is often hard to disentangle from the volatility of the shocks that move the same expectations: a lower ρ and a smaller shock variance both compress the forecast distribution (Remark 8). The local diagnostics show no severe confound of this form, though they do not establish global identification (Online Appendix Table [OA.6](#) and Figure [OA.4](#)). Only one volatility parameter is materially correlated with ρ in the posterior, that of the consumption-preference shock (+0.50; the next largest correlation with a shock or measurement-error standard deviation is -0.20), and all of them together explain 27% of the posterior variance of ρ . In two-dimensional likelihood slices, letting the preference-shock volatility adjust raises the standard error of ρ by only 3%, and letting its persistence adjust by 9%. Attention is estimated jointly with the demand block: its largest correlations are with the Taylor-rule response to the output gap (-0.55), the preference-shock volatility, habit (+0.31) and the inverse Frisch elasticity (+0.26). Fixing all shock and measurement-error standard deviations would lower the standard error of ρ only from 0.023 to 0.018; fixing those four demand-block parameters would lower it to 0.010. Treating a variable uniformly across equations disciplines ρ , but it does not guarantee that ρ cannot absorb other misspecification.

5.7.4 Discounting variables or discounting terms?

An alternative way to bring cognitive discounting into a nonlinear model, used by [Hirose et al. \(2024\)](#) and implicit in [Gabaix \(2020\)](#), discounts the expectation of each forward-looking *term* toward its steady-state value, $\mathbb{E}_t^{\text{BR}} g(X_{t+1}) - \bar{g} = M (\mathbb{E}_t g(X_{t+1}) - \bar{g})$, rather than each perceived *variable*. Section 2 implies that, with equal attentions and matching reference values, the two rules coincide at first order and differ beyond it. The term-wise rule shrinks the expected deviation of the term by M , so the risk correction inside that expectation is weighted by M . The operator instead scales the one-step k -th cumulant of each variable's

chart deviation by the k -th power of its attention (Proposition 3). With a common attention, it therefore discounts risk by ρ^2 . We check both statements on the estimated model (Table 14, Figure 12). The comparison is conducted under the common restriction, for tractability: a single attention matches the term-wise version of the laboratory model one-to-one to a single M . It discounts every stationary forward term of every equation except the exogenous preference shifter and the habit argument of the consumption condition (the literal version below discounts those too).

At first order the two rules are the same: with $M = \rho$ at the estimated parameters the log-likelihoods differ by 1.5×10^{-11} and the impulse responses by 10^{-15} . A literal version that also discounts the exogenous preference shifter and the habit term inside the consumption condition differs already at first order, by 0.45 log points. Which terms are discounted is thus a modeling choice under the term-wise rule, the ambiguity of Gap 2. With one coefficient and coherent reference values the rule is linear in each term, so the difference comes from the choice of terms, not from how they are written. Beyond first order the rules differ. Under the term-wise rule, the precautionary shift of inflation (risky minus deterministic steady state) is -0.048 percentage points a year, against -0.024 under the operator. The second-order part of the inflation response to a three-standard-deviation preference shock is 9.7% of the response under the term-wise rule, against 0.4% under the operator, and of the opposite sign. Both are small in absolute terms at the estimated parameters. The comparison is not a clean test of ρ^2 against M : technology enters the operator through its moving reference point and is not discounted, whereas the term-wise rule scales the whole term. Along a fixed-parameter likelihood slice, with the other parameters at the common-restriction mode, the second-order likelihood peaks at the same attention under both rules (0.895); this is a slice, not a profile likelihood, since the other parameters are not re-optimized. At those parameters the operator's second-order likelihood is 13 log points above the first-order-matched term-wise rule and the literal variant is 0.4 log points above the operator; these fixed-parameter comparisons do not establish a model ranking. The difference lies in what must be chosen. The operator also requires choices: which variables to discount, the chart, the reference point. But they are made once per variable and hold in every equation, whereas the term-wise rule must be specified term by term.

5.7.5 The lower bound

The sample includes the 2009–2015 lower-bound episode, and every estimate confronts it with the occasionally binding constraint $R_t = \max\{\underline{r}, R_t^*\}$ of Section 3. Table 12 shows that the bound has almost no effect on the estimates of the common restriction: dropping it changes the log-likelihood by less than one log point in either model, leaves ρ at 0.896, and

Table 14: Operator versus term-wise discounting under the common restriction, at the estimated parameters ($\rho = M = 0.897$). Panel A: first order; panel B: second-order likelihood, and the peak of a likelihood slice in attention with the other parameters fixed at the mode and not re-optimized (the row “Peak along the fixed-parameter slice”); panel C: risky minus deterministic steady state; panel D: second-order correction to a +3 standard-deviation preference shock. The rational column uses the rational model at its own mode. “Term-wise (literal)” also discounts the exogenous preference shifter and the habit term; its standard error is from the curvature at the bounded-search optimum.

	RE	Operator	Term-wise	Term-wise (literal)
<i>A. First order (validation)</i>				
log L , CD mode (RE: RE mode)	4234.1658	4244.0459	4244.0459	4243.59
difference from operator			-1.5×10^{-11}	-0.45
log L , RE mode, attention = 1			4234.1658	4234.1658
<i>B. Second-order likelihood</i>				
log L , CD mode (RE: RE mode)	4132.17	4171.22	4158.24	4171.58
difference from operator			-12.98	+0.36
Peak along the fixed-parameter slice		0.895	0.895	0.897
curvature s.e.		0.006	0.003	0.009
log L at peak		4171.26	4158.30	4171.58
<i>C. Risky minus deterministic steady state</i>				
Inflation (p.p. ann.)	-0.2255	-0.0241	-0.0480	-0.0644
Shadow policy rate (p.p. ann.)	-0.3840	-0.0497	-0.0796	-0.1028
Output growth (p.p. ann.)	-0.0000	+0.0000	+0.0000	-0.0000
Output gap (%)	-0.0343	-0.0158	-0.0158	-0.0180
Consumption, detrended (%)	-0.0354	-0.0158	-0.0159	-0.0181
<i>D. Order-2 correction to a +3 s.d. preference shock: peak (share of first-order peak)</i>				
Inflation (p.p. ann.)	+0.0090 (7.0%)	+0.0008 (0.4%)	-0.0190 (9.7%)	-0.0282 (11.5%)
Shadow policy rate (p.p. ann.)	+0.0111 (0.8%)	+0.0023 (0.2%)	-0.0237 (2.0%)	-0.0335 (2.3%)
Output gap (%)	+0.0246 (1.4%)	+0.0035 (0.3%)	-0.0089 (0.8%)	-0.0107 (0.8%)

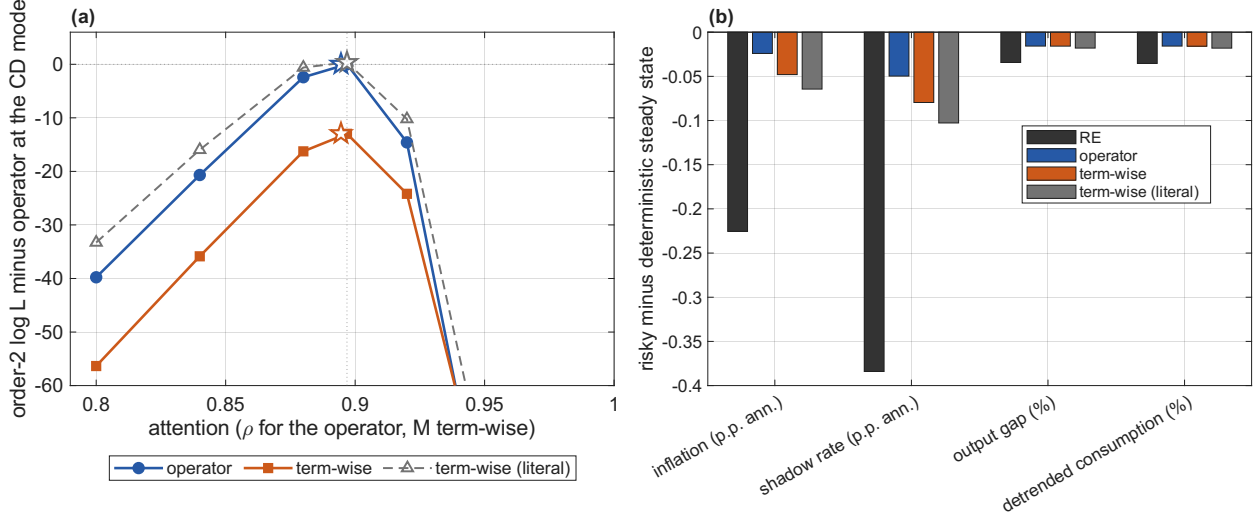


Figure 12: Operator versus term-wise discounting under the common restriction. (a) Second-order log-likelihood relative to the operator at the mode, as a function of attention (ρ for the operator, M for the term-wise rule), other parameters fixed at the mode (a slice); stars mark the peaks along the slices. (b) Risky minus deterministic steady state.

leaves the margin in favor of cognitive discounting essentially unchanged (14.5, computed from unrounded values). Under the common restriction, then, the ranking is similar when the bound is omitted, under the reported implementation. This is a sensitivity check for that specification; it does not test the accuracy of the constrained likelihood or the variable-specific comparisons. Attention is the same in and out of the bound; letting it differ at the bound is a question the operator can pose (Proposition 4) and that we take up, at the calibration, in Online Appendix [OA-F.2](#).

5.8 How the estimates compare with existing ones

The posterior means of the four attentions the data inform most, between 0.76 and 0.91 per quarter, sit toward the upper end of the range in the cognitive-discounting literature, around the value Gabaix calibrates. [Gabaix \(2020\)](#) calibrates attention at 0.85; [Hirose et al. \(2024\)](#), estimating a small fully nonlinear New Keynesian model with the lower bound, find that households and firms discount the future by around 15% ($M \approx 0.85$). In linearized estimated models, [Meggiorini \(2023\)](#) finds that the data prefer a substantial degree of discounting. [Afsar et al. \(2024\)](#), who discipline the discount by the under-revision of survey forecasts in a hybrid model with backward-looking households and firms, find values between about 0.2 and 0.4 per quarter. [Hohberger et al. \(2024\)](#) find a strong preference for cognitive discounting ($m \approx 0.4$ for all agents) and, when the Phillips curve gets its own parameter, price setters close to rational, as our sector-specific estimates do. [Meggiorini and Milani \(2026\)](#) show that

the estimate depends on how expectations are otherwise modeled: about 0.4 under rational expectations, about 0.93–0.94 under Euler-equation learning.

Two observations discipline the comparison. First, at first order and with equal attentions the operator coincides with the device estimated in those studies, up to the IS-curve term of Remark 2 (Proposition 1); what differs is the model, the observables and the econometrics. The present model is comparatively rich — habit, sticky wages and prices, a credit spread, a stochastic trend, seven observables with estimated measurement errors, and a lower bound — and each of these features plausibly supplies persistence and expectational damping that a smaller model must otherwise purchase with myopia. Second, per-quarter attentions of this size are less moderate than they look: through the forward solution, news about the distant future is discounted geometrically (Proposition 1), which is what, at given structural parameters, moves the forward-guidance results of §4.4 away from rational expectations.

5.9 A methodological note

The estimates rest on four ingredients (Online Appendix OA-E). Measurement errors are estimated, which lets the likelihood attribute part of each observation to noise; whether any quarters dominate the likelihood would require observation-level contributions, which are not reported here. The posterior mode is located by a global search, which makes a poor local optimum less likely but does not certify a global maximum. Because the rational model is nested, the behavioral model’s maximized likelihood is at least the rational one. This concerns maximized likelihoods, not the posterior density at the mode or the marginal likelihood, whose priors and dimensions differ. At the reported modes the behavioral log-likelihood exceeds the rational one. The diffuse component of the state is initialized exactly. And marginal likelihoods are checked across estimators. Empirical support for or against bounded rationality is only as good as these ingredients, and it is conditional on the approximate filter.

6 Conclusion

A variable-specific behavioral-expectations operator applies cognitive discounting consistently in nonlinear, growing and regime-switching models. It is defined directly on the nonlinear equilibrium conditions, with explicit choices of variables, coordinates and reference points. At second order and above, it also discounts *risk*. It complements formulations that contract the perceived state or discount selected expected terms. At first order the operator reproduces Gabaix’s (2020) one-step discounting of expectations, except that each selected

variable is discounted in every equation in which it appears (Remark 2). At second order it scales the loading on each conditional variance and covariance of future variables' deviations from a fixed reference point by the product of their attentions (by the square of a common attention; Proposition 2). More generally, it scales each variable's k -th one-step conditional cumulant of the chart deviation by the k -th power of its attention (Proposition 3). Under the conditions of Corollary 2, it therefore turns divergent conditional expectations into convergent ones, a candidate structural complement to pruning. Finally, it lets attention itself switch with the state of the economy (Proposition 4).

In the laboratory model, the economic lessons of linear cognitive discounting carry over, with thresholds set mainly by the model's structure: cognitive discounting relaxes the Taylor principle, weakens forward guidance away from the lower bound at given structural parameters, and raises the government-spending multiplier — the results of Gabaix (2020), here inside a nonlinear, nonstationary, estimated model. Two results go beyond the linear device. First, in the laboratory model's simulations, low enough attention removes the explosions of higher-order dynamics that are driven by persistence, the channel the operator acts on. Lower attention is needed as persistence rises. No value on our grid suffices at the highest persistence. Explosions driven by large shocks are not removed in general, and at second order the operator can also create explosions at intermediate attention. It complements pruning rather than replacing it. Second, with matching common attention and reference values, discounting variables and discounting terms coincide at first order but treat risk differently. A rule that discounts each forward-looking term weights the risk correction inside it by the discount. The operator instead weights the second-order loading by the square of the common attention, or by products of attentions in general. It also scales the k -th one-step cumulant of each variable's chart deviation by the k -th power of its attention. With common attention, the operator therefore discounts risk more than the mean (Proposition 2, Remark 4). In the laboratory model with a common attention, a permanent peg at the occasionally binding bound is locally determinate only below a threshold $\rho^* \in [0.720, 0.736]$. Above it, the effect of calendar forward guidance at the bound keeps growing with the length of the peg over the horizons computed. With attention by variable, determinacy depends on the whole attention vector.

Because the operator is defined on the estimable model, the framework can put the behavioral question to the data. We estimate the nonstationary model with an occasionally binding lower bound on pre-2020 US data. With attention estimated variable by variable, the data decisively prefer cognitive discounting to rational expectations, by about 14 log points of marginal likelihood. The margin is at least 12.8 log points under every attention structure estimated. Posterior mean attention is uneven across variables: it is lowest, with wide

posteriors, on the reference inflation rates, and the data give positive evidence for freeing those two attentions jointly rather than imposing rational expectations on both. Attention by variable is positively favored over block-specific and common restrictions, while a households/firms grouping fits about as well in this model — a result learned by comparing the restrictions with the full variable-by-variable model. The estimation illustrates the method in one model with one set of observables: in another model, or with other observables such as survey expectations, the right grouping may differ, which is why the operator is built variable by variable. It is designed for richer models, where restrictions such as a common or block-specific discount become testable rather than assumed.

The operator is implemented in RISE, so the object analyzed in the theory and the object solved and estimated in the applications are the same.

References

- Afsar, A., J.-E. Gallegos, R. Jaimes, and E. Silgado-Gómez (2024). “A Behavioral Hybrid New Keynesian Model: Quantifying the Importance of Belief Formation Frictions.” *Economic Modelling* 132, 106626.
- Andrade, J., P. Cordeiro, and G. Lambais (2019). “Estimating a Behavioral New Keynesian Model.” arXiv:1912.07601.
- Andreasen, M.M., J. Fernández-Villaverde, and J.F. Rubio-Ramírez (2018). “The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications.” *Review of Economic Studies* 85(1), 1–49.
- Angeletos, G.-M., and C. Lian (2018). “Forward Guidance without Common Knowledge.” *American Economic Review* 108(9), 2477–2512.
- Angeletos, G.-M., and Z. Huo (2021). “Myopia and Anchoring.” *American Economic Review* 111(4), 1166–1200.
- Angeletos, G.-M., and C. Lian (2023). “Determinacy without the Taylor Principle.” *Journal of Political Economy* 131(8), 2125–2164.
- Aruoba, S.B., P. Cuba-Borda, and F. Schorfheide (2018). “Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries.” *Review of Economic Studies* 85(1), 87–118.
- Ascari, G., S. Mavroeidis, and N. McClung (2023). “Coherence without Rationality at the Zero Lower Bound.” *Journal of Economic Theory* 214, 105745.
- Benchimol, J., and L. Bounader (2023). “Optimal Monetary Policy Under Bounded Rationality.” *Journal of Financial Stability* 67, 101151.
- Bianchi, F. (2013). “Regime Switches, Agents’ Beliefs, and Post-World War II U.S. Macroeconomic Dynamics.” *Review of Economic Studies* 80(2), 463–490.
- Binning, A., and J. Maih (2015). “Sigma Point Filters for Dynamic Nonlinear Regime Switching Models.” Norges Bank Working Paper 10/2015.
- Binning, A., and J. Maih (2017). “Modelling Occasionally Binding Constraints Using Regime-Switching.” Norges Bank Working Paper 23/2017.
- Bordalo, P., N. Gennaioli, and A. Shleifer (2018). “Diagnostic Expectations and Credit Cycles.” *Journal of Finance* 73(1), 199–227.

- Bracha, A., and J. Tang (2025). “Inflation Levels and (In)Attention.” *Review of Economic Studies* 92(3), 1564–1594.
- Chib, S., and I. Jeliazkov (2001). “Marginal Likelihood from the Metropolis–Hastings Output.” *Journal of the American Statistical Association* 96(453), 270–281.
- Coibion, O., and Y. Gorodnichenko (2015). “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts.” *American Economic Review* 105(8), 2644–2678.
- Davig, T., and E.M. Leeper (2007). “Generalizing the Taylor Principle.” *American Economic Review* 97(3), 607–635.
- Del Negro, M., M.P. Giannoni, and C. Patterson (2023). “The Forward Guidance Puzzle.” *Journal of Political Economy Macroeconomics* 1(1), 43–79.
- Den Haan, W.J., and J. De Wind (2012). “Nonlinear and Stable Perturbation-Based Approximations.” *Journal of Economic Dynamics and Control* 36(10), 1477–1497.
- Durbin, J., and S.J. Koopman (2012). *Time Series Analysis by State Space Methods*, 2nd ed. Oxford: Oxford University Press.
- Evans, G.W., and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.
- Farhi, E., and I. Werning (2019). “Monetary Policy, Bounded Rationality, and Incomplete Markets.” *American Economic Review* 109(11), 3887–3928.
- Fernández-Villaverde, J., G. Gordon, P. Guerrón-Quintana, and J.F. Rubio-Ramírez (2015). “Nonlinear Adventures at the Zero Lower Bound.” *Journal of Economic Dynamics and Control* 57, 182–204.
- Foerster, A., J.F. Rubio-Ramírez, D.F. Waggoner, and T. Zha (2016). “Perturbation Methods for Markov-Switching Dynamic Stochastic General Equilibrium Models.” *Quantitative Economics* 7(2), 637–669.
- Gabaix, X. (2014). “A Sparsity-Based Model of Bounded Rationality.” *Quarterly Journal of Economics* 129(4), 1661–1710.
- Gabaix, X. (2016a). “Behavioral Macroeconomics via Sparse Dynamic Programming.” NBER Working Paper 21848.

- Gabaix, X. (2016b). “A Behavioral New Keynesian Model.” NBER Working Paper 22954 (revised 2019).
- Gabaix, X. (2020). “A Behavioral New Keynesian Model.” *American Economic Review* 110(8), 2271–2327.
- García-Schmidt, M., and M. Woodford (2019). “Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis.” *American Economic Review* 109(1), 86–120.
- Gelfand, A.E., and D.K. Dey (1994). “Bayesian Model Choice: Asymptotics and Exact Calculations.” *Journal of the Royal Statistical Society, Series B* 56(3), 501–514.
- Geweke, J. (1999). “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication.” *Econometric Reviews* 18(1), 1–73.
- Hamilton, J.D. (1989). “A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle.” *Econometrica* 57(2), 357–384.
- Hirose, Y., H. Iiboshi, M. Shintani, and K. Ueda (2024). “Estimating a Behavioral New Keynesian Model with the Zero Lower Bound.” *Journal of Money, Credit and Banking* 56(8), 2185–2197.
- Hohberger, S., A. Ifrim, B. Pataracchia, and M. Ratto (2024). “Myopic Behaviour in Macroeconomic Models: Empirical Evidence from the US.” CAMA Working Paper 11/2024.
- Holden, T.D. (2023). “Existence and Uniqueness of Solutions to Dynamic Models with Occasionally Binding Constraints.” *Review of Economics and Statistics* 105(6), 1481–1499.
- Hommes, C., K. Mavromatis, T. Özden, and M. Zhu (2023). “Behavioral Learning Equilibria in New Keynesian Models.” *Quantitative Economics* 14(4), 1401–1445.
- Ilabaca, F., G. Meggiorini, and F. Milani (2020). “Bounded Rationality, Monetary Policy, and Macroeconomic Stability.” *Economics Letters* 186, 108522.
- Karaboga, D., and B. Basturk (2007). “A Powerful and Efficient Algorithm for Numerical Function Optimization: Artificial Bee Colony (ABC) Algorithm.” *Journal of Global Optimization* 39(3), 459–471.
- Kass, R.E., and A.E. Raftery (1995). “Bayes Factors.” *Journal of the American Statistical Association* 90(430), 773–795.

- Kim, J., S. Kim, E. Schaumburg, and C.A. Sims (2008). “Calculating and Using Second-Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models.” *Journal of Economic Dynamics and Control* 32(11), 3397–3414.
- Kolasa, M., S. Ravgotra, and P. Zabczyk (2025). “Monetary Policy and Exchange Rate Dynamics in a Behavioral Open Economy Model.” *Journal of International Economics* 155, 104087.
- Koopman, S.J., and J. Durbin (2003). “Filtering and Smoothing of State Vector for Diffuse State-Space Models.” *Journal of Time Series Analysis* 24(1), 85–98.
- L’Huillier, J.-P., S.R. Singh, and D. Yoo (2024). “Incorporating Diagnostic Expectations into the New Keynesian Framework.” *Review of Economic Studies* 91(5), 3013–3046.
- Liu, Z., D.F. Waggoner, and T. Zha (2011). “Sources of Macroeconomic Fluctuations: A Regime-Switching DSGE Approach.” *Quantitative Economics* 2(2), 251–301.
- Lombardo, G., and H. Uhlig (2018). “A Theory of Pruning.” *International Economic Review* 59(4), 1825–1836.
- Lubik, T.A., and M. Marzo (2021). “Fiscal Policy Perceptions in a Behavioral New Keynesian Model.” *Annals of Economics and Finance* 22(2), 255–287.
- Maćkowiak, B., and M. Wiederholt (2009). “Optimal Sticky Prices under Rational Inattention.” *American Economic Review* 99(3), 769–803.
- Maih, J. (2010). “Conditional Forecasts in DSGE Models.” Norges Bank Working Paper 07/2010.
- Maih, J. (2015). “Efficient Perturbation Methods for Solving Regime-Switching DSGE Models.” Norges Bank Working Paper 01/2015.
- McKay, A., E. Nakamura, and J. Steinsson (2016). “The Power of Forward Guidance Revisited.” *American Economic Review* 106(10), 3133–3158.
- McKay, A., E. Nakamura, and J. Steinsson (2017). “The Discounted Euler Equation: A Note.” *Economica* 84(336), 820–831.
- Meggiorini, G. (2023). “Behavioral New Keynesian Models: An Empirical Assessment.” *Journal of Macroeconomics* 77, 103538.

- Meggiorini, G., and F. Milani (2026). “Behavioral New Keynesian Models: Learning versus Cognitive Discounting.” *Journal of Money, Credit and Banking*, early view, doi:10.1111/jmcb.70086.
- Meng, X.-L., and W.H. Wong (1996). “Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration.” *Statistica Sinica* 6, 831–860.
- Nakata, T. (2017). “Optimal Government Spending at the Zero Lower Bound: A Non-Ricardian Analysis.” *Review of Economic Dynamics* 23, 150–169.
- Pfäuti, O. (2025). “Inflation—Who Cares? Monetary Policy in Times of Low Attention.” *Journal of Money, Credit and Banking* 57(5), 1211–1239.
- Pfäuti, O., and F. Seyrich (2022). “A Behavioral Heterogeneous Agent New Keynesian Model.” CRC TR 224 Discussion Paper 334.
- Sims, C.A. (2003). “Implications of Rational Inattention.” *Journal of Monetary Economics* 50(3), 665–690.
- Sims, C.A., D.F. Waggoner, and T. Zha (2008). “Methods for Inference in Large Multiple-Equation Markov-Switching Models.” *Journal of Econometrics* 146(2), 255–274.
- Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner (2021). “Rank-Normalization, Folding, and Localization: An Improved $\hat{R}$ for Assessing Convergence of MCMC.” *Bayesian Analysis* 16(2), 667–718.
- Weber, M., B. Candia, H. Afrouzi, T. Ropele, R. Lluberas, S. Frache, B. Meyer, S. Kumar, Y. Gorodnichenko, D. Georgarakos, O. Coibion, G. Kenny, and J. Ponce (2025). “Tell Me Something I Don’t Already Know: Learning in Low- and High-Inflation Settings.” *Econometrica* 93(1), 229–264.
- Woodford, M. (2019). “Monetary Policy Analysis when Planning Horizons are Finite.” *NBER Macroeconomics Annual* 33, 1–50.

A The laboratory model

This appendix gives the laboratory model of Section 3 exactly as it is coded and solved: its variables, its 31 equilibrium conditions, the treatment of the lower bound, the cognitive-discounting operator, the steady state and the calibration. With it, the model can be rebuilt and run from the paper alone. The equations are numbered in the order of the model file. Online Appendix OA-G derives every condition from the agents' problems and writes out the equations of each variant used in the paper. For the stationary policy variant it writes out the private conditions and the general form of the planner's conditions.

Notation. Time- t variables are written without the expectation operator; every term dated $t + 1$ is inside $\mathbb{E}_t$ under rational expectations and inside $\mathbb{E}_t^{\text{BR}}$ under cognitive discounting. Gross rates are $\Pi_t \equiv 1 + \pi_t$, $\Pi_t^w \equiv 1 + \pi_t^w$, $\tilde{\Pi}_t \equiv 1 + \tilde{\pi}_t$ and $\tilde{\Pi}_t^w \equiv 1 + \tilde{\pi}_t^w$. A subscript ss denotes the steady-state value of a variable, computed from the model rather than set as a parameter; for a trending variable it is the value on the balanced growth path at the same date. A subscripted ρ (ρ_r , $\rho_{\pi h}$, $\rho_{\pi w}$, ρ_s , ρ_g , ρ_θ , $\rho_{\theta w}$, $\rho_{\zeta c}$, $\rho_{\zeta h}$) is an interest-rate smoothing, indexation or autoregressive coefficient, not an attention. Attention is ρ_X with a variable X as subscript (ρ_C , ρ_Λ , ρ_P , ρ_Y , ρ_π , ρ_{π^w} , $\rho_{\tilde{\pi}}$, $\rho_{\tilde{\pi}^w}$), ρ_H , ρ_F , ρ_{Euler} in the restrictions of §5.5, and the unsubscripted ρ in the restriction of §5.7. Likewise, Λ_t and the spread loading χ_s are unrelated to the stakes $\bar{\Lambda}$ of Online Appendix OA-A.5 and to the chart χ of Section 2, and κ is the Phillips-curve slope of the three-equation model (with an index, $\kappa_{k,t}$ is a cumulant and κ_i an attention cost).

Variables and shocks. Table 15 lists the 31 endogenous variables and the 15 shocks. The eight trending variables grow along the balanced growth path and are solved in logarithms; the others are stationary. All shocks are independent standard normal.

Equilibrium conditions. The household has preferences $\mathbb{E}_0 \sum_t \beta^t [\zeta_{C,t} \log(C_t - bC_{t-1}) - \zeta_{H,t} H_t^{1+\nu} / (1 + \nu)]$ with internal habit b , sets wages subject to a Rotemberg cost, and holds a bond whose return carries the credit spread. Firms produce with $Y_t = A_t H_t$, finance the wage bill at the spread, and set prices subject to a Rotemberg cost. The shifter d multiplies the discount factor; it equals one in the occasionally binding model and differs from one only in the floor regime of the two-regime version below.

Table 15: Endogenous variables and shocks of the laboratory model.

<i>Trending (solved in logs)</i>		<i>Stationary</i>	
A_t	technology	$\zeta_{C,t}$	consumption-preference shifter
Λ_t	marginal utility of nominal wealth	$\zeta_{H,t}$	labor-disutility shifter
C_t	consumption	θ_t	goods elasticity (price markup)
Y_t	output	θ_t^w	labor elasticity (wage markup)
W_t	nominal wage	g_t	government-spending share
P_t	price level	H_t	hours
G_t	government spending	π_t, π_t^w	net price and wage inflation
D_t	firm profits	$\tilde{\pi}_t, \tilde{\pi}_t^w$	reference price and wage inflation
		S_t	real marginal cost
		R_t, R_t^*	policy rate, shadow rate
		$\tilde{y}_t$	output gap
		Δy_t	log output growth
		s_t	credit spread
<i>Observables:</i> $R_t^{obs}, \pi_t^{obs}, s_t^{obs}, \Delta y_t^{obs}, \Delta C_t^{obs}, \Delta W_t^{obs}, \Delta g_t^{obs}$.			
<i>Structural shocks:</i> $\varepsilon_{\zeta c,t}, \varepsilon_{\zeta h,t}, \varepsilon_{\theta,t}, \varepsilon_{\theta w,t}, \varepsilon_{A,t}, \varepsilon_{g,t}, \varepsilon_{R,t}, \varepsilon_{s,t}$.			
<i>Measurement errors:</i> $e_t^R, e_t^\pi, e_t^s, e_t^y, e_t^c, e_t^w, e_t^g$.			

Preference shifters.

$$\log \zeta_{C,t} = (1 - \rho_{\zeta c}) \log \bar{\zeta}_C + \rho_{\zeta c} \log \zeta_{C,t-1} + \sigma_{\zeta c} \varepsilon_{\zeta c,t}, \quad (\text{A.1})$$

$$\log \zeta_{H,t} = (1 - \rho_{\zeta h}) \log \bar{\zeta}_H + \rho_{\zeta h} \log \zeta_{H,t-1} + \sigma_{\zeta h} \varepsilon_{\zeta h,t}. \quad (\text{A.2})$$

Resource constraint: output is consumed by households and the government or absorbed by the two adjustment costs,

$$Y_t = C_t + G_t + \frac{\phi_p}{2\bar{\Pi}_t} \left(\frac{\Pi_t}{\bar{\Pi}_t} - 1 \right)^2 Y_t + \frac{\phi_w}{2} \left(\frac{\Pi_t^w}{\bar{\Pi}_t^w} - 1 \right)^2 Y_t. \quad (\text{A.3})$$

Reference wage inflation (partial indexation to last period's wage inflation):

$$\tilde{\pi}_t^w = (1 - \rho_{\pi w}) \tilde{\pi}_{ss}^w + \rho_{\pi w} \pi_{t-1}^w. \quad (\text{A.4})$$

Consumption first-order condition (the second term is the effect of current consumption on next period's habit) and *Euler equation* (the spread raises the return on saving):

$$\Lambda_t P_t = \frac{\zeta_{C,t}}{C_t - bC_{t-1}} - \beta db \frac{\zeta_{C,t+1}}{C_{t+1} - bC_t}, \quad (\text{A.5})$$

$$\Lambda_t = \beta d (1 + R_t) (1 + k_s s_t) \Lambda_{t+1}. \quad (\text{A.6})$$

Wage Phillips curve (wage setting with a Rotemberg cost and indexation):

$$0 = \theta_t^w \zeta_{H,t} H_t^\nu - (\theta_t^w - 1) \Lambda_t W_t - \phi_w \frac{\Lambda_t P_t Y_t \Pi_t^w}{H_t \tilde{\Pi}_t^w} \left(\frac{\Pi_t^w}{\tilde{\Pi}_t^w} - 1 \right) + \beta d \phi_w \frac{\Lambda_{t+1} P_{t+1} Y_{t+1} \Pi_{t+1}^w}{H_t \tilde{\Pi}_{t+1}^w} \left(\frac{\Pi_{t+1}^w}{\tilde{\Pi}_{t+1}^w} - 1 \right). \quad (\text{A.7})$$

Markup shifters and technology (a unit root with drift):

$$\log \theta_t^w = (1 - \rho_{\theta w}) \log \bar{\theta}^w + \rho_{\theta w} \log \theta_{t-1}^w + \sigma_{\theta w} \varepsilon_{\theta w, t}, \quad (\text{A.8})$$

$$\log \theta_t = (1 - \rho_\theta) \log \bar{\theta} + \rho_\theta \log \theta_{t-1} + \sigma_\theta \varepsilon_{\theta, t}, \quad (\text{A.9})$$

$$\log A_t = \log(1 + \mu_a) + \log A_{t-1} + \sigma_A \varepsilon_{A, t}. \quad (\text{A.10})$$

Real marginal cost (the wage bill is financed at the spread) and *reference price inflation*:

$$S_t = \frac{W_t}{P_t A_t} (1 + k_s s_t), \quad (\text{A.11})$$

$$\tilde{\pi}_t = (1 - \rho_{\pi h}) \pi^* + \rho_{\pi h} \pi_{t-1}. \quad (\text{A.12})$$

Price Phillips curve, the symmetric first-order condition of price setting with the nominal cost $\frac{\phi_p}{2\tilde{\Pi}_t} \left(\frac{p_{i,t}}{p_{i,t-1}\tilde{\Pi}_t} - 1 \right)^2 P_t Y_t$; the term $\chi_s s_t \theta_t$ lets the spread shift prices directly (it is zero at the calibration):

$$0 = (1 - \theta_t) + \theta_t S_t + \chi_s s_t \theta_t - \phi_p \frac{\Pi_t}{\tilde{\Pi}_t^2} \left(\frac{\Pi_t}{\tilde{\Pi}_t} - 1 \right) + \beta d \phi_p \frac{\Lambda_{t+1} Y_{t+1} \Pi_{t+1}^2}{\Lambda_t Y_t \tilde{\Pi}_{t+1}^2} \left(\frac{\Pi_{t+1}}{\tilde{\Pi}_{t+1}} - 1 \right). \quad (\text{A.13})$$

Firm profits (before the financing of the wage bill, which is paid to households):

$$D_t = P_t Y_t - W_t H_t - \frac{\phi_p}{2\tilde{\Pi}_t} \left(\frac{\Pi_t}{\tilde{\Pi}_t} - 1 \right)^2 P_t Y_t. \quad (\text{A.14})$$

Government spending:

$$\log g_t = (1 - \rho_g) \log \bar{g} + \rho_g \log g_{t-1} + \sigma_g \varepsilon_{g, t}, \quad (\text{A.15})$$

$$G_t = g_t Y_t. \quad (\text{A.16})$$

Monetary policy: a Taylor rule with smoothing sets the shadow rate, and the policy rate is

the shadow rate truncated at the lower bound $\underline{r}$:

$$\begin{aligned} \log(1 + R_t^*) &= \log(1 + R_{ss}^*) + \rho_r [\log(1 + R_{t-1}^*) - \log(1 + R_{ss}^*)] \\ &\quad + (1 - \rho_r) \phi_\pi [\log \Pi_t - \log(1 + \pi^*)] + (1 - \rho_r) \phi_y \tilde{y}_t + \sigma_r \varepsilon_{R,t}, \end{aligned} \quad (\text{A.17})$$

$$R_t = \max\{\underline{r}, R_t^*\}. \quad (\text{A.18})$$

Output gap, inflation identities, production and output growth:

$$\tilde{y}_t = \log Y_t - \log A_t - \log H_{ss}, \quad (\text{A.19})$$

$$\Pi_t = P_t / P_{t-1}, \quad (\text{A.20})$$

$$\Pi_t^w = W_t / W_{t-1}, \quad (\text{A.21})$$

$$Y_t = A_t H_t, \quad (\text{A.22})$$

$$\Delta y_t = \log(Y_t / Y_{t-1}). \quad (\text{A.23})$$

Credit spread:

$$s_t = \rho_s s_{t-1} + \sigma_s \varepsilon_{s,t}. \quad (\text{A.24})$$

Observation equations (the measurement errors are independent; their standard deviations are estimated):

$$R_t^{obs} = R_t + \sigma_{me,R} e_t^R, \quad (\text{A.25})$$

$$\pi_t^{obs} = \pi_t - \pi^* + \sigma_{me,\pi} e_t^\pi, \quad (\text{A.26})$$

$$s_t^{obs} = s_t + \sigma_{me,s} e_t^s, \quad (\text{A.27})$$

$$\Delta y_t^{obs} = \log(Y_t / Y_{t-1}) - \log(1 + \mu_a) + \sigma_{me,y} e_t^y, \quad (\text{A.28})$$

$$\Delta c_t^{obs} = \log(C_t / C_{t-1}) - \log(1 + \mu_a) + \sigma_{me,c} e_t^c, \quad (\text{A.29})$$

$$\Delta w_t^{obs} = \log(W_t / W_{t-1}) - \log \Pi_t - \log(1 + \mu_a) + \sigma_{me,w} e_t^w, \quad (\text{A.30})$$

$$\Delta g_t^{obs} = \log(G_t / G_{t-1}) - \log(1 + \mu_a) + \sigma_{me,g} e_t^g. \quad (\text{A.31})$$

Equations (A.1)–(A.31) are 31 conditions for the 31 endogenous variables.

The lower bound. Three versions of (A.18) appear in the paper.

- *Occasionally binding:* $R_t = \max\{\underline{r}, R_t^*\}$ with $d = 1$. It is used for every estimate with the bound (Section 5), including the block-specific and term-wise comparisons, and for every lower-bound result of Section 4: the peg-determinacy threshold (§4.1), calendar forward guidance at the bound (§4.4), the cost of the bound (§4.5) and the multiplier

at the bound (§4.6).

- *No bound*: $R_t = R_t^*$ with $d = 1$. It is used for the higher-order stability results (§4.2), the impulse responses and moments (§4.3), forward guidance away from the bound (§4.4), and the estimates without the bound.
- *Two-regime Markov bound*: a Markov chain $z_t \in \{1, 2\}$ (normal, floor) with constant transition probabilities p_{12} and p_{21} replaces (A.18) by

$$R_t = a_{z_t} R_t^* + (1 - a_{z_t})(\underline{r} + \sigma_{r,z_t} \varepsilon_{R,t}), \quad (\text{A.18}')$$

with $(a, d, \sigma_r) = (1, 1, \sigma_{r,1})$ in the normal regime and $(0, d_{\text{floor}}, \sigma_{r,2})$ at the floor; the policy-shock standard deviation in (A.17) is regime-specific as well. The floor value $d_{\text{floor}} = (1 + \pi^*)(1 + \mu_a)/[\beta(1 + \underline{r})]$ makes the floor regime's steady state consistent with $R = \underline{r}$ (Binning and Maih, 2017). Attention may differ across the two regimes. Its floor regime is computed in Online Appendix OA-F.2, whose threshold §4.1 and §4.4 quote; with its two regimes made identical it is the version without the bound, and the stability results of §4.2 are computed in that form.

The optimal-policy results of §4.7 use a stationary variant without the bound (Online Appendix OA-G.13).

Cognitive discounting. The operator of Definition 1 is applied to the eight forward-looking endogenous variables $\{C, \Lambda, P, Y, \pi, \pi^w, \tilde{\pi}, \tilde{\pi}^w\}$, which exhausts every endogenous variable dated $t + 1$ in (A.5), (A.6), (A.7) and (A.13). The exogenous preference shifter $\zeta_{C,t+1}$ is not discounted. Each discounted variable is contracted in the same way in every equation in which it appears:

- the four rates, in the level chart toward their steady states: $x_{t+1} \mapsto x_{ss} + \rho_x(x_{t+1} - x_{ss})$ for $x \in \{\pi, \pi^w, \tilde{\pi}, \tilde{\pi}^w\}$;
- the four levels, in the log chart toward a reference point that moves with their trends: $X_{t+1} \mapsto \bar{X}_{t+1} (X_{t+1}/\bar{X}_{t+1})^{\rho_x}$, with

$$\begin{aligned} \bar{C}_{t+1} &= C_{ss,t+1} \frac{A_{t+1}}{A_{ss,t+1}}, & \bar{Y}_{t+1} &= Y_{ss,t+1} \frac{A_{t+1}}{A_{ss,t+1}}, \\ \bar{P}_{t+1} &= (1 + \pi^*)P_t, & \bar{\Lambda}_{t+1} &= \Lambda_{ss,t+1} \frac{A_{ss,t+1}}{A_{t+1}} \frac{P_{ss,t}}{P_t}. \end{aligned} \quad (\text{A.32})$$

The references use realized technology for real quantities and the lagged price level grown at trend inflation for the price level; Λ , whose trend is that of $1/(AP)$, uses the inverse of

both. For consumption and output, discounting the level is then discounting it relative to realized technology; for the price level the contracted object is gross inflation, and for Λ it is $\Lambda_{t+1}A_{t+1}P_t$. Because the anchor is random, the risk results of Section 2 apply to the detrended deviations, conditional on the anchor: risk in the anchor itself (technology) is not discounted (§2.5; Online Appendix OA-C.1). The “perceived household” reading of the operator is local, as discussed in §2.2. In the baseline estimation each of the eight variables has its own attention ρ_X (§5.4); one attention ρ for all eight is the robustness restriction of §5.7. The operator leaves the steady state unchanged, because every contraction is the identity on the balanced growth path.

Steady state. On the balanced growth path, technology, output, consumption and government spending grow at the gross rate $g_z \equiv 1 + \mu_a$, the price level at $1 + \pi^*$, the nominal wage and profits at $(1 + \pi^*)g_z$, and Λ at $1/[(1 + \pi^*)g_z]$; technology and the price level are normalized to one at the base date. The stationary variables and the base-date levels are, in this order,

$$\begin{aligned} \zeta_C &= \bar{\zeta}_C, \quad \zeta_H = \bar{\zeta}_H, \quad \theta = \bar{\theta}, \quad \theta^w = \bar{\theta}^w, \quad g = \bar{g}, \quad s = 0, \\ \pi &= \tilde{\pi} = \pi^*, \quad \pi^w = \tilde{\pi}^w = (1 + \pi^*)g_z - 1, \quad S = \frac{\bar{\theta} - 1}{\bar{\theta}}, \quad W = S P A, \\ 1 + R^* &= \frac{(1 + \pi^*)g_z}{\beta}, \quad R = R^*, \quad H = \left[\frac{(\bar{\theta}^w - 1) \bar{\zeta}_C (g_z - \beta d b) (\bar{\theta} - 1)}{(g_z - b) (1 - \bar{g}) \bar{\theta}^w \bar{\theta} \bar{\zeta}_H} \right]^{1/(1+\nu)}, \\ Y &= A H, \quad C = (1 - \bar{g})Y, \quad G = \bar{g}Y, \quad \Lambda = \frac{\bar{\zeta}_C (g_z - \beta d b)}{(g_z - b) C P}, \\ D &= \frac{P Y}{\bar{\theta}}, \quad \tilde{y} = 0, \quad \Delta y = \log g_z, \end{aligned}$$

and the observables are $R^{obs} = R$ and zero for the other six. With $d = 1$ the steady state requires $R_{ss}^* > \underline{r}$, which holds at the calibration. In the floor regime of the two-regime version, $d = d_{\text{floor}}$ and $R = \underline{r}$; the formulas above, which carry βd , then give the floor regime’s own steady state, with lower hours, output, consumption, government spending, profits and Λ , while R^* keeps the value above. The labor-disutility level $\bar{\zeta}_H$ is a calibrated parameter, chosen so that steady-state hours are about one third; H is solved from the wage Phillips curve as above.

Calibration. Table 16 lists the parameter values used at the calibration. The estimated parameters and their priors are in Table 17; the posterior estimates are in Section 5 and Online Appendix OA-E.

Table 16: Calibration of the laboratory model.

Parameter	Description	Value
<i>Preferences and technology</i>		
β	discount factor	0.99
ν	inverse Frisch elasticity	2
b	habit	0.70
$\bar{\theta}, \bar{\theta}^w$	elasticities among goods, among labor types	6
$\bar{\zeta}_C$	consumption-preference level	1
$\bar{\zeta}_H$	labor-disutility level	24.7066
$\bar{g}$	government-spending share	0.20
μ_a	trend growth (quarterly)	0.0065
<i>Nominal rigidities and policy</i>		
ϕ_p, ϕ_w	price, wage adjustment costs	58.25, 96.15
$\rho_{\pi h}, \rho_{\pi w}$	price, wage indexation	0.50
ϕ_π, ϕ_y	Taylor rule: inflation, output gap	1.50, 0.50
ρ_r	interest-rate smoothing	0.75
π^*	inflation target (quarterly)	0.005
$\underline{r}$	lower bound (quarterly)	0.000348
σ_r	policy shock	0.0015
<i>Shocks</i>		
$\rho_g, \rho_\theta, \rho_{\theta w}, \rho_{\zeta c}, \rho_{\zeta h}$	persistences	0.85
σ_A, σ_g	technology, government spending	0.007, 0.008
$\sigma_\theta, \sigma_{\theta w}, \sigma_{\zeta c}, \sigma_{\zeta h}$	markups, preferences	0.005
<i>Credit spread</i>		
ρ_s, σ_s	persistence, innovation	0.95, 0.0005
k_s	loading in the Euler equation and marginal cost	1
χ_s	loading in the price Phillips curve	0
<i>Two-regime Markov bound</i>		
p_{12}, p_{21}	transition probabilities	0.025, 0.040
$\sigma_{r,1}, \sigma_{r,2}$	policy shock, normal and floor regimes	0.0015, 0.000197

The adjustment costs are the Rotemberg equivalents of Calvo probabilities of not resetting of 0.75 for prices and 0.80 for wages, $\phi = (\bar{\theta} - 1)\omega / [(1 - \omega)(1 - \beta\omega)]$. Trend growth is the mean of quarterly log real GDP growth over the full data set, 1985Q1–2026Q1 (165 quarters); the lower bound and the floor-regime shock standard deviation are the mean and standard deviation of the policy rate in the 39 quarters of that data set in which it is below 0.1 percent. The estimation sample ends in 2019Q4. At the calibration the measurement-error standard deviations are 10^{-4} ; in the estimation they are estimated. All attentions equal one at the calibration (rational expectations) unless stated otherwise.

Variants. Online Appendix [OA-G](#) gives the complete equations of the other models used in the paper: the two-regime Markov bound, the block-specific and term-wise versions of §5.5 and §5.7.4, the stationary policy variant of §4.7, and the detrended twin used to check the levels model.

Prior distributions

Table 17 lists the priors. Parameters not listed (the discount factor, the adjustment costs, steady-state targets, the spread block, and the policy-shock standard deviation) are fixed at their calibrated values.

Table 17: Prior distributions.

Parameter	Description	Distribution	Mean	Std. dev.
ρ_X , eight X	attention, variable by variable	Uniform	on $[0, 1]$	
b	habit	Beta	0.70	0.10
ν	inverse Frisch elasticity	Gamma	2.00	0.50
ϕ_π	Taylor rule, inflation	Gamma	1.50	0.25
ϕ_y	Taylor rule, output gap	Gamma	0.50	0.20
ρ_r	interest-rate smoothing	Beta	0.75	0.10
$\rho_{\pi h}, \rho_{\pi w}$	price, wage indexation	Beta	0.50	0.15
$\rho_g, \rho_\theta, \rho_{\theta w}, \rho_{\zeta c}, \rho_{\zeta h}$	persistences	Beta	0.75	0.12
σ_A	technology shock	Inv. gamma	0.007	0.004
σ_g	government-spending shock	Inv. gamma	0.008	0.004
$\sigma_\theta, \sigma_{\theta w}, \sigma_{\zeta c}, \sigma_{\zeta h}$	markup, preference	Inv. gamma	0.005	0.003
σ_{me}	measurement errors	Inv. gamma	$0.4 \hat{s}$	$0.4 \hat{s}$
<i>Robustness restriction (§5.7)</i>				
ρ	common attention	Beta	0.85	0.10

$\hat{s}$: approximate sample standard deviation of the observed series. In the sector- and block-specific versions of §5.5 each free attention parameter has the Uniform prior on $[0, 1]$; in their Beta-prior estimates (Online Appendix [OA-E.2](#)) it has the prior of ρ .

B Where to find the details

Each formal statement has one maintained source. Results stated in Section 2 are proved in the Online Appendix, in sections titled “Proof of ...” that list the assumptions first; the perceived-law results are summarized in §2.7 and stated in full, with their proofs, only in the Online Appendix, where they are numbered by section (Proposition [OA-A.1](#), ...). Table 18 is the map.

Table 18: Where to find each result, its proof, and the technical details.

Result or material	Main text	Proof or details
Definition 1 (the operator); mean action (7); joint feasibility	§2.2	formal setting: §OA-A.2
Proposition 1 (first order: Gabaix)	§2.3	proof: §OA-A.1
Proposition 2 (discounting of risk)	§2.4	proof: §OA-B.1
Corollary 1 (precautionary wedge)	§2.4	proof and non-Gaussian example: §OA-B.2; numerical check: §OA-B.5
Proposition 3 (cumulant contraction)	§2.5	proof: §OA-B.3
Corollary 2 (finite behavioral expectations)	§2.5	proof, and the threshold itself: §OA-B.4
Moving reference point for trending variables	§2.6	§OA-C.1
Proposition 4 (regime switching)	§2.6	proof: §OA-C.2
Perceived-law representation	§2.7	Proposition OA-A.1, §OA-A.2
Horizons and composed forecasts	§2.7	Proposition OA-A.2, §OA-A.3
Model consistency; state versus variable contraction	§2.7	Proposition OA-A.3, Corollary OA-A.1, §OA-A.4
Sparse-max reading and its scope	§2.7	Proposition OA-A.4, §OA-A.5
Three-equation determinacy criteria	§4.1	§OA-D.1
Three-equation forward matrix	§4.4	§OA-D.2
Laboratory model: equations, steady state, calibration, priors	Section 3	Appendix A; derivations: §OA-G
Model variants: attention switching at the bound, block-specific, term-wise, stationary policy variant, detrended twin	§§5.5, 5.7.4, 4.7	§§OA-G.9–OA-G.13
Estimation: likelihood, optimizer, sampler	Section 5	§OA-E
Attention variable by variable (baseline): identification, diagnostics	§§5.4–5.6	§OA-E.1
Common-attention restriction (robustness): posterior, priors, identification, sector and block under Beta priors	§5.7	§OA-E.2
Cost of the lower bound: full numbers	§4.5	§OA-F.1
The lower bound as a Markov regime	§4.1, §4.4	§OA-F.2