

Optimal Monetary Policy

When the Economy Must Be Learned

Frank Schmid

This version: 31 October 2026

First version: 6 September 2026

Abstract. Optimal monetary policy at the frontier is loss-function policy, optimal against a model. This paper asks what optimal policy is when the model must be learned from the record, and finds that learning enters at four levels, each measured on real-time data for the United States. The economy's agents learn, yet three certified constructions of the expectations block select model-consistent expectations for the agents in the macro rows. The econometrician learns the model, whose likelihood does not pin the scale of the policy block, which a registered convention fixes. A correction learns what the structure misses, and its evidence of signal must be judged against a null that carries a floor spell. The policymaker learns the rule: a policy trained by reinforcement inside the estimated economy reproduces the liftoff timing of the optimal-control path and sits above the Federal Open Market Committee's (FOMC) realized decisions throughout the record. That distance is learned from the record: an agent frozen after the first floor spell, whose estimate does not identify the steeper Phillips curve, stays near the FOMC, while an agent re-learned at every decision point departs from it once the record contains the inflation of 2021.

JEL classification: E52, E58, E37, C45, D83.

This paper was developed with assistance from generative AI tools, including Claude, across the design, analysis, coding, and drafting stages. The author supervised all stages, verified the results, and takes full responsibility for the content.

Contents

Abbreviations and Acronyms	ii
1 Introduction	1
2 The Real-Time Record	3
3 Optimal Policy, Rules, and Learners	5
3.1 The book's experiment on the learned rule	6
4 The World Model	7
4.1 Blocks and expectations	8
4.2 Estimator and certification	8
4.3 The multi-start search and the estimate of record	9
4.4 The model's forecasts against the staff's	11
5 The Expectations Block Against Three Families	12
6 Residual Components and a Floor-Aware Null	15
7 The Policy Agent	16
7.1 Loss and training environment	16
7.2 Structured or learned?	16
7.3 Training streams and the checkpoint of record	17
8 The Optimal Policy and the FOMC	19
8.1 The agent of record against the FOMC	19
8.2 The agent frozen at the end of the first floor spell	22
8.3 Counterfactual outcomes	27
9 Sensitivity and Limitations	29
10 Conclusion	32
A The Oracle Conditions	33
Bibliography	34

Abbreviations and Acronyms

ACM	Adrian–Crump–Moench term-premium model
ALFRED	Archival Federal Reserve Economic Data
CBO	Congressional Budget Office
CPI	consumer price index
DM1	dual-mandate loss function of Adrian, Laxton, and Obstfeld, chapter 12, at its baseline weights (DM2 halves the weight on the output gap)
ELB	effective lower bound
FOMC	Federal Open Market Committee
FRB/US	the Federal Reserve Board’s large-scale model of the United States economy
FRED	Federal Reserve Economic Data
GP	Gaussian process
HLW	Holston–Laubach–Williams neutral-rate model
IFB	inflation-forecast-based (rule)
IS	investment–saving (block)
LOO	leave-one-out
MAP	maximum a posteriori
NW	Newey–West (standard error)
PCE	personal consumption expenditures
RMS	root mean square
RMSE	root-mean-square error
RTDSM	Real-Time Data Set for Macroeconomists
SE	standard error
SPF	Survey of Professional Forecasters

Greek letters

α (alpha)	weight on squared inflation deviations in the loss
β (beta)	weight on the squared output gap in the loss
γ (gamma)	weight on the squared policy-rate change in the loss
η (eta)	innovation to the policy-rule shock
κ (kappa)	slope of the Phillips block
λ_1 (lambda)	weight on expected four-quarter inflation in the Phillips block
π (pi)	inflation
ρ (rho)	interest-rate smoothing in the policy rule
σ (sigma)	standard deviation
ϕ_u (phi)	persistence of the policy-rule shock

1 Introduction

The frontier of monetary policy, as Adrian, Laxton, and Obstfeld describe it, is policy that minimizes an explicit loss over the forecast rather than policy that follows a simple rule.¹ The loss is quadratic in the deviation of inflation from target, the output gap, and the change in the policy rate; because it is quadratic, the response it implies is increasingly strong as deviations grow, which is the risk-avoidance property the book argues a central bank needs near a dark corner such as the effective lower bound (ELB); the phrase is Blanchard's, the book credits it to him where it introduces the loss, and the working paper the book's model comes from is titled for it.² In the book's own experiment from the state of 2014:Q4, the loss-minimizing path holds the policy rate at the floor into 2016 while an inflation-forecast-based (IFB) rule lifts off at once. That optimal policy is optimal against a calibrated model, and the book's confidence bands come from the history of forecast errors around it.

This paper asks what optimal policy is when the economy must be learned from the record. The question is not rhetorical: a policy that minimizes a loss inside a model is only as good as the model, and it can be worse, because a learner that optimizes against a model exploits whatever the model gets wrong. Learning enters at four levels, and the paper measures each. The *economy's agents* learn, or do not: the expectations block is where forward guidance at a floor has its effects and where the behavioral and learning literatures have proposed departures from model-consistent expectations.³ The *econometrician* learns the model, by an estimator whose behavior under a floor spell is certified on a synthetic oracle before it touches the record, and by a search that must find the likelihood's optimum rather than a basin near its starting point. A *correction* learns what the structure misses, a pair of nonparametric components on the model's residuals, whose evidence of signal must be judged against a null that carries a floor spell. And the *policymaker* learns the rule, by reinforcement inside the estimated economy, with the exploitation the method invites bounded by the estimate's own uncertainty and its checkpoint selected by a criterion from a search over training streams. Figure 1 draws the four learners and the check bound to each.

The world model is a reduced form of the book's semi-structural economy, an investment–saving (IS) relation, a hybrid Phillips curve, an expectations block that is this paper's, and a policy slot with an ELB, estimated on 122 quarters of real-time data from 1996:Q1 through 2026:Q2, a record that carries 36 quarters at the floor in two spells and the inflation of 2021–2023. The policy agent is an IFB rule with learned coefficients trained on the real-time states of every quarter and compared with the Federal Open Market Committee's (FOMC) realized decisions at those states, with bands from the estimate's uncertainty.

Four findings organize the paper. First, the economy's agents. The long-run survey expectation follows a constant-gain learner's anchor at a measured speed, but the agents in the macro rows do not lean on it: three constructions of the expectations block, each nesting model-consistent expectations at one parameter value and each certified under a floor spell, select that value on the record, with the likelihood profile in the alternative's parameter steep and monotone. Second, the

¹Tobias Adrian, Douglas Laxton, and Maurice Obstfeld, eds., *Advancing the Frontiers of Monetary Policy* (Washington, DC: International Monetary Fund, 2018), chap. 12.

²Olivier Blanchard, "Where Danger Lurks," *Finance & Development* 51, no. 3 (2014); Ali Alich, Kevin Clinton, Charles Freedman, Michel Juillard, Ondra Kamenik, Douglas Laxton, Jarkko Turunen, and Hou Wang, "Avoiding Dark Corners: A Robust Monetary Policy Framework for the United States," IMF Working Paper 15/134 (Washington, DC: International Monetary Fund, 2015).

³Xavier Gabaix, "A Behavioral New Keynesian Model," *American Economic Review* 110, no. 8 (2020): 2271–2327; George W. Evans and Seppo Honkapohja, *Learning and Expectations in Macroeconomics* (Princeton, NJ: Princeton University Press, 2001).

econometrician. The likelihood does not pin the scale of the policy block, which the paper fixes by a convention registered before any estimate under it; under the convention the deepest basin, about 120 units of log-likelihood above the estimate the data-driven start reaches, is found from the record's own catalogued points and not by a random search, and it carries a resolved Phillips slope, a nearly forward-looking Phillips block, and a persistent policy shock. Third, the correction. A floor spell puts structure into the filtered shocks at the true parameters, so the certification statistic's null shifts by an order of magnitude, from a 95th percentile of 0.079 on pre-floor economies to 0.385 on economies with a floor; against the floor-aware null the components carry no signal in the inflation block, and an agent trained on them beats the textbook rule on fewer than half of the draws from the estimate's uncertainty, where the agent of record beats it on every draw. Fourth, the policymaker. Trained on the estimated blocks under the estimate's uncertainty and selected by a validation criterion from a search over training streams, the optimal rule under the book's loss is an inertial rule with an inflation response of 2.17 and a gap response of 2.18, whose prescriptions track the first-difference rule of the FOMC's own reporting more closely than the Taylor (1993) rule. From the book's 2014:Q4 state it holds the floor for five quarters and lifts off in 2016:Q2, one quarter after the book's optimal-control path and five after the IFB rule; it beats the textbook rule in its own world on 100 of 100 draws from the estimate's uncertainty; and across every environment since 1997 it sits +0.768 points above the FOMC's realized rate on average, outside the band the estimate's uncertainty draws on 88 of 118 quarters, with the first floor spell the exception, where it holds the floor as the FOMC did. Frozen at 2015:Q4, with the estimate, the components, and the training confined to the record before liftoff, the same construction prescribes rates +0.235 points above the FOMC on average over the 42 quarters that follow, against the agent of record's +1.297 on the same quarters: the record before liftoff does not identify the Phillips slope that the inflation of 2021–2023 identifies, and the distance of the optimal rule from the FOMC is largely learned from that episode.

The paper stands on four literatures. Real-time policy evaluation: Orphanides shows that a rule evaluated on revised data misstates what a policymaker could have done with the data of the day, which is why every state the agent sees here is a vintage.⁴ Optimal control at the Federal Reserve: the loss-minimizing paths of the book are of the construction the Board's staff run in the FRB/US model, and the simple rules the FOMC reports in its Monetary Policy Report, the Taylor (1993) rule among them, are the comparators of Section 8.1.⁵ Policy under model uncertainty: Brainard's attenuation, the robust-control program of Hansen and Sargent, and the parameter-uncertainty results of Wieland and of Edge, Laubach, and Williams describe how a policymaker who knows the model is uncertain acts, and Sargent and Primiceri give the policymaker's own learning a history; the agent here is trained under the estimate's uncertainty, which is the reinforcement-learning form of the same concern.⁶ Reinforcement learning in monetary economics: Hinterlang

⁴Athanasios Orphanides, "Monetary Policy Rules Based on Real-Time Data," *American Economic Review* 91, no. 4 (2001): 964–985.

⁵Flint Brayton, Thomas Laubach, and David Reifschneider, "Optimal-Control Monetary Policy in the FRB/US Model," FEDS Notes (Washington, DC: Board of Governors of the Federal Reserve System, November 21, 2014); John B. Taylor, "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39 (1993): 195–214; John B. Taylor and John C. Williams, "Simple and Robust Rules for Monetary Policy," in *Handbook of Monetary Economics*, vol. 3B, ed. Benjamin M. Friedman and Michael Woodford (Amsterdam: Elsevier, 2010), 829–859.

⁶William C. Brainard, "Uncertainty and the Effectiveness of Policy," *American Economic Review* 57, no. 2 (1967): 411–425; Lars Peter Hansen and Thomas J. Sargent, *Robustness* (Princeton, NJ: Princeton University Press, 2008); Volker Wieland, "Monetary Policy, Parameter Uncertainty and Optimal Learning," *Journal of Monetary Economics* 46, no. 1 (2000): 199–228; Rochelle M. Edge, Thomas Laubach, and John C. Williams, "Welfare-Maximizing Monetary Policy under Parameter Uncertainty," *Journal of Applied Econometrics* 25, no. 1 (2010): 129–143; Thomas J. Sargent, *The Conquest of American Inflation* (Princeton, NJ: Princeton University Press, 1999); Giorgio E. Primiceri, "Why

and Tänzer train a policy by reinforcement against a neural transition model estimated on United States data, and Chen and coauthors use deep reinforcement learning as a model of bounded rationality inside a general-equilibrium model; the construction here, a policy trained inside a learned world model, follows the Dyna architecture of Sutton and the world-model program of Ha and Schmidhuber.⁷ What distinguishes this paper from each is stated where it arises: the record is real-time throughout, the estimator is certified on an oracle under a floor spell, the expectations block is tested against learning and behavioral alternatives, and the policy's exploitation of the learned model is bounded and measured.

The paper proceeds as follows. Section 2 describes the real-time record. Section 3 states the concept of optimal policy the book sets out, the distinction between an optimal path and an optimal rule, and the four learners, and Section 3.1 runs the book's experiment on the learned rule. Section 4 states the world model, its estimator, the oracle certification, the search that fixes the estimate of record, and the estimate's forecasts against the staff's. Section 5 tests the expectations block against the three families. Section 6 fits the residual components and constructs the floor-aware null. Section 7 trains the agent, decomposes what an agent trained with the components exploits, and selects the rule of record by a search over training streams. Section 8 compares the agent of record with the FOMC, repeats the comparison with everything frozen at the end of the first floor spell, and simulates the outcomes each rule would have produced with the record's own shocks. Section 9 reports the sensitivity to the loss weights and to the training stream and states the limitations of the construction, and Section 10 concludes. Appendix A lists the oracle conditions.

2 The Real-Time Record

Every observation the model sees at a decision point is the observation the FOMC could have seen. The record is built from a vintage store of 30 series holding every release of each series as its own payload: the Archival Federal Reserve Economic Data (ALFRED) vintages of the core personal consumption expenditures (PCE) price index, the unemployment rate, the effective funds rate, and the two-year Treasury yield among others; the Philadelphia Fed's Real-Time Data Set for Macroeconomists (RTDSM) and its Tealbook staff projections of the output gap and core inflation; the Survey of Professional Forecasters (SPF) at the four-quarter and ten-year horizons; the Holston–Laubach–Williams (HLW) neutral-rate estimates;⁸ the Adrian–Crump–Moench (ACM) term premiums;⁹ and the Wu–Xia shadow rate.¹⁰ The decision points are the final scheduled FOMC meeting of each quarter from 1990:Q1 through 2026:Q2, 146 points, each assembled from the latest

Inflation Rose and Fell: Policy-Makers' Beliefs and U.S. Postwar Stabilization Policy," *Quarterly Journal of Economics* 121, no. 3 (2006): 867–901.

⁷Natascha Hinterlang and Alina Tänzer, "Optimal Monetary Policy Using Reinforcement Learning," Discussion Paper 51/2021 (Frankfurt am Main: Deutsche Bundesbank, 2021); Mingli Chen, Rama Cont, Andreas Joseph, Michael Kumhof, Xinlei Pan, Wei Xiong, and Xuan Zhou, "Deep Reinforcement Learning in a Monetary Model," Staff Working Paper 1,142 (London: Bank of England, 2025); Richard S. Sutton, "Dyna, an Integrated Architecture for Learning, Planning, and Reacting," *ACM SIGART Bulletin* 2, no. 4 (1991): 160–163; David Ha and Jürgen Schmidhuber, "Recurrent World Models Facilitate Policy Evolution," in *Advances in Neural Information Processing Systems* 31 (2018): 2450–2462.

⁸Kathryn Holston, Thomas Laubach, and John C. Williams, "Measuring the Natural Rate of Interest: International Trends and Determinants," *Journal of International Economics* 108 (2017): S59–S75.

⁹Tobias Adrian, Richard K. Crump, and Emanuel Moench, "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics* 110, no. 1 (2013): 110–138.

¹⁰Jing Cynthia Wu and Fan Dora Xia, "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound," *Journal of Money, Credit and Banking* 48, no. 2–3 (2016): 253–291.

vintage of each series released strictly before the meeting; the store was rebuilt from source on 5 September 2026. The filter also reads the moments of the record's measurement errors, and each is computed from what was known at the decision point it serves: the staff gap's revision, nowcast, and projection errors and their correlations, over the quarters whose fifth-later estimate existed by then (112 quarters at 2026:Q2); the survey's four-quarter expected-inflation error, from 2000:Q1 on the same rule (87 quarters); and the term premium through which the two-year rate is measured, as quarterly means of the ACM two-year premium over the quarters completed before the decision point (121 quarters, 1996:Q1 through 2026:Q1, from the release the store holds, retrieved 5 September 2026). The premium is measured without a convenience component, and the window closes on years in which the dollar's convenience yield moved: Krishnamurthy and Chumbo find the convenience yield on short-term dollar safe assets about where it stood in 2019 by 2025 and that on ten-year Treasuries fallen over the same years; the record reads the two-year premium alone.¹¹ The estimate of record and the agents trained on it read the moments known at 2026:Q2, the frozen agent of Section 8.2 those known at its cutoff, and each rolling window those known at its decision point. One more set, the error of the real-time HLW estimate against its ex post value (32 vintages at 2026:Q2, the first from 2016), enters only the simulated economies of the oracle of Section 4.2 and no estimate on the record.

Computing the moments from what was known at each decision point replaced moments measured once on samples ending in 2019:Q4. The change was registered before any estimate under it was computed, with five predictions of its own: three on the rolling agent, reported in Section 8.2, and two reported here. Both are read on the estimate and the agents this paper reports. The estimate and the agents also carry four later changes, each of which alters the estimate, the agents, or both: the convention on the policy block's inflation response, the multi-start search (Section 4.3), the search over training streams, and the redrawn validation starts (Section 7.3); two others, a profile of the objective and the reading of the selection on a fresh set, which retrains nothing, alter neither. Each reading below was fixed after its outcome could be seen. The first prediction was that the expectations block would move from the earlier estimate's by more than half "a standard error of the record," with λ_1 remaining at its structural limit. The record is read as the earlier estimate, and the move is tested weight by weight; remaining at the limit is read as the estimator's flag for λ_1 equal to one, the definition written into the registered code for the rolling agent's prediction on λ_1 (Section 8.2) before any outcome. The prediction held in its first part and failed in its second, and so failed: a_{lag} and a_{lead} moved by 0.079 in opposite directions, against half their earlier standard errors, 0.011 and 0.013 (half the new estimate's would give the same verdict, as would the larger of each pair), and λ_1 , which the earlier estimate placed at the limit, stands at 0.946 with a standard error of 0.041 (Section 4.3). The two weights' sum is held at the cap of Section 4.3, so the two moves are one. The second prediction was that the frozen agent's coefficients would remain a fraction of the agent of record's, as in the comparison the registration cited. The coefficients are read as the inflation and gap responses, and a fraction as a ratio below one; the reading on all four coefficients is set aside because on it the prediction fails in that comparison itself. It failed: the frozen agent's responses stand at 1.23 and 1.34 times the agent of record's (Section 8.2), and the failure does not rest on which stream either search selected: every one of the frozen search's eight streams has larger responses than every one of the agent of record's search's eight streams, by factors above 1.03 in the inflation response and above 1.09 in the gap response. Of the other readings considered (for the first prediction: the new estimate's standard errors, and λ_1 within one or two

¹¹Arvind Krishnamurthy and Miguel Chumbo, "Disappearing Dollar Convenience?" NBER Working Paper 35742 (Cambridge, MA: National Bureau of Economic Research, 2026).

standard errors of the limit; for the second: all four coefficients, the long-run responses, and ratios no larger than in the cited comparison), only one would change a verdict: with λ_1 within two standard errors of the limit, a test the registration did not state and one resting on a standard error computed near the limit, the first prediction would hold.

Under the change alone, before the later changes, the first verdict was the same, on narrower margins in both parts: the two weights moved by 0.030 and λ_1 stood at 0.995, so the first prediction failed by the flag; λ_1 was within one of its standard errors of the limit, and on that reading it would have held there. The second prediction held under the change alone, the frozen agent's responses at 0.61 and 0.30 times the agent of record's, and under each later change until the validation starts were redrawn, the ratios rising toward one with each change and standing at 0.85 and 0.95 times just before the redraw. It failed when they were redrawn: the redrawn starts lowered the agent of record's responses, and the frozen search, re-run under them, kept each stream's checkpoint of record and the selection. This history holds on the reading adopted here; on all four coefficients the prediction never held under any version, and on the long-run responses and the cited comparison's ratios it failed under earlier changes as well.

Three constructions carry the record past the floor. The staff projections, released with a five-year lag, are admitted at a decision point only from a vintage dated within that point's own quarter, so the final 22 quarters carry no staff rows and the inflation, funds-rate, two-year-rate, and survey rows carry the sample there. The floor is observed rather than inferred: a quarter is on the floor where the target range's lower bound at the decision point is zero, which gives 36 floor quarters in two spells, 2009:Q1–2015:Q4 and 2020:Q2–2022:Q1; at such a quarter the rate row is censored at the model's floor. The neutral-rate path is the ex post one-sided HLW estimate on the window from its latest vintage. The estimation window is 1996:Q1 through 2026:Q2, 122 quarters and 836 observation rows.

3 Optimal Policy, Rules, and Learners

The book's optimal policy minimizes, at each decision, the loss

$$L_t = \sum_{i \geq 1} [\alpha (\pi_{t+i}^4 - \pi^*)^2 + \beta y_{t+i}^2 + \gamma (i_{t+i} - i_{t+i-1})^2], \quad (1)$$

with π^4 four-quarter inflation, π^* the target, y the output gap, i the policy rate, and weights α and β of one and γ of one half; the book's sum runs from the current quarter, $i \geq 0$, and the sum here from one quarter ahead. The quadratic form is the point: a large deviation weighs disproportionately, so the response to a shock is stronger the larger the shock and the nearer the economy stands to a corner from which recovery is hard, and the book's simulations show the loss-minimizing policy holding the floor for a lengthy period after a negative shock where a linear rule lifts off early. The book contrasts this policy with two rules, the Taylor rule and the IFB rule, which respond linearly to a small set of variables and carry no such risk avoidance.

Two distinctions organize what follows. The first is between an optimal *path* and an optimal *rule*. The book's DM1 policy, the label the book gives its dual-mandate loss at these weights, is a path: the sequence of rates that minimizes equation (1) from the current state under the model, re-optimized as the state moves. The agent of this paper is a rule: an IFB form with four coefficients chosen to minimize the same loss on average over the states the record supplies, inside the estimated economy. A rule is what a central bank can commit to and what a comparison with realized

decisions can be made against; a path is the benchmark the rule is measured by, and Section 3.1 measures the rule against it on the book's own experiment.

The second distinction is between the four senses in which something is learned, drawn in Figure 1. The economy's agents learn, in the sense of the expectations literature: whether their expectation of inflation is the model's forecast, an anchor, a discounted forecast, or a level they update from experience. The econometrician learns the blocks from the record, in the statistical sense, and the paper certifies that learner on an oracle and characterizes the likelihood it searches. The residual correction learns nonparametrically what the blocks miss, and the paper tests whether what it learns is signal. And the policymaker learns the rule by reinforcement, which is learning by optimization against a model that is itself learned, so that the first three learners' errors are the fourth learner's opportunity. The paper's claim about optimal policy is conditional on all four, and it states at each step what would change the reading.

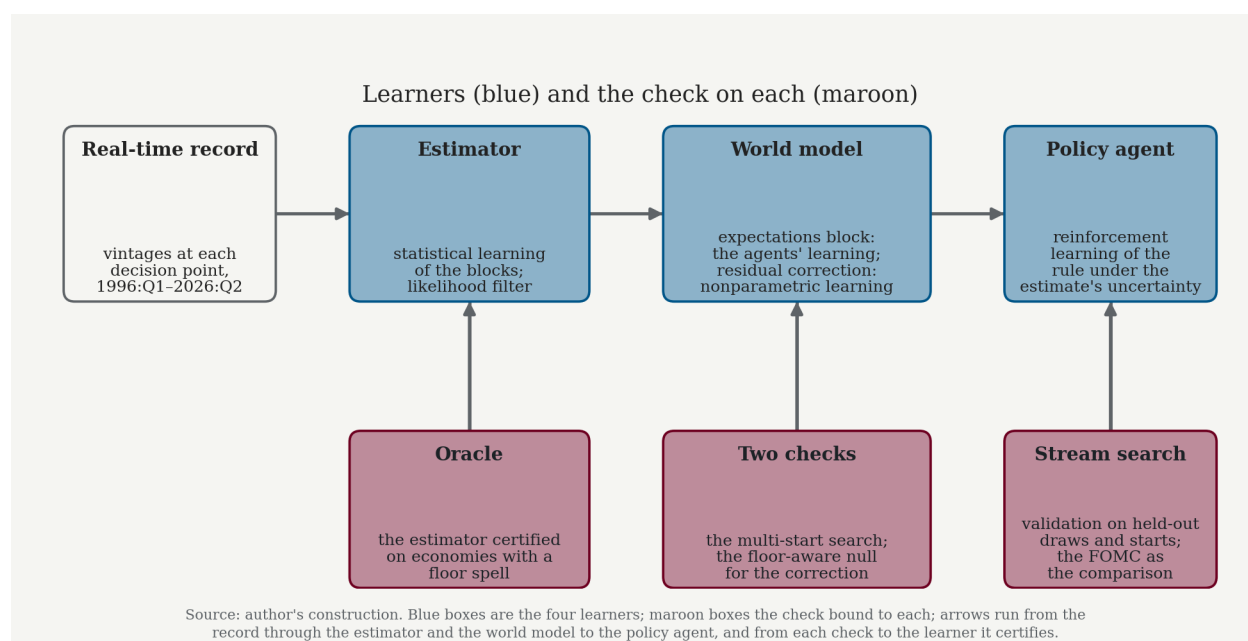


Figure 1: The four learners of optimal monetary policy when the economy must be learned. The top row runs from the real-time record through the estimator, which learns the blocks, and the world model, whose expectations block carries the economy's agents' learning and whose residual correction learns what the blocks miss, to the policy agent, which learns the rule by reinforcement under the estimate's uncertainty. The bottom row is the check bound to each learner: the oracle certification of the estimator on economies with a floor spell, the multi-start search and the floor-aware null for the world model, and the validation search over training streams and the comparison with the FOMC for the agent. Blue boxes are learners and maroon boxes checks. Source: author's construction.

3.1 The book's experiment on the learned rule

The book's Figure 12.5 starts the economy at its 2014:Q4 state, a gap of -2 percent, four-quarter inflation of 1.5 percent, and the rate at the floor, and runs the IFB rule and the loss-minimizing DM1 path forward with no shocks. Figure 2 runs the same experiment in the book economy on

the learned rule of Section 7.3, the agent of record, with DM1 re-optimized each quarter under one construction. The IFB rule lifts off in 2015:Q1; DM1 holds the floor and lifts off in 2016:Q1; the learned rule holds the floor for five quarters and lifts off in 2016:Q2. The losses are 14.44 for the IFB rule, 9.95 for DM1, and 10.64 for the learned rule. A rule learned by reinforcement inside an economy estimated from the record reproduces, to within a quarter, the liftoff timing of the book's optimal-control path, which is the risk-avoidance property the book argues for and the simple rule lacks; the rule's loss lies between the path's and the simple rule's, as a rule constrained to four coefficients must.

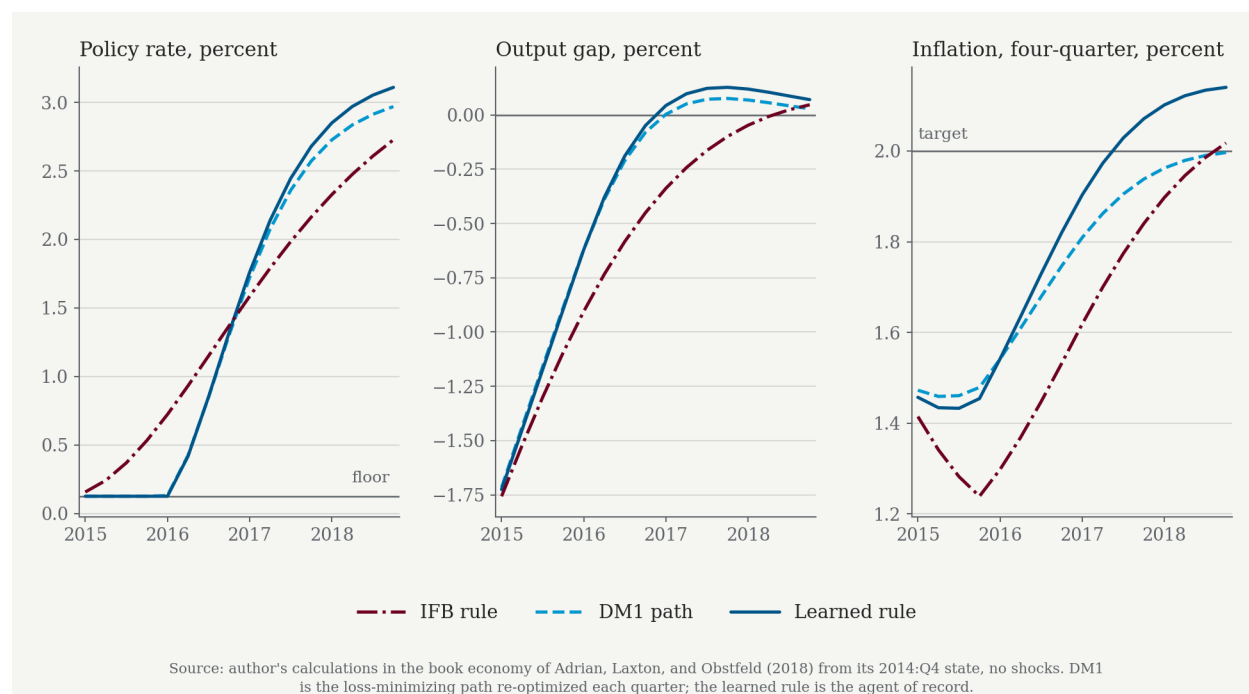


Figure 2: The book's experiment from the 2014:Q4 state run on the learned rule. Each panel carries one variable over 2015:Q1 to 2018:Q4: the policy rate, the output gap, and four-quarter inflation, all in percent, under the IFB rule at the book coefficients (maroon, dash-dot), the loss-minimizing DM1 path re-optimized each quarter (bright blue, dashed), and the agent of record (deep blue, solid), in the book economy with no shocks; gray lines mark the floor, zero, and the target. Liftoff: the IFB rule in 2015:Q1, DM1 in 2016:Q1, the learned rule in 2016:Q2; losses 14.44, 9.95, and 10.64. Source: author's calculations in the book economy of Adrian, Laxton, and Obstfeld.

4 The World Model

The world model is the semi-structural economy of Adrian, Laxton, and Obstfeld, estimated on the record of Section 2 by a certified likelihood filter; this section states its blocks, its estimator and the estimator's certification, the search that fixes its estimate of record, and the estimate's forecasts against the staff's.

4.1 Blocks and expectations

The economy is this paper's reduced form of the semi-structural model of Adrian, Laxton, and Obstfeld, written here as it is solved. Chapter 12 prints the model's Phillips curve and its policy rules and describes its output-gap equation in words alone; that equation, its calibration, and the rest of the model are in the working paper the chapter builds on, a stripped-down Global Projection Model.¹² The IS block is

$$y_t = a_{\text{lag}} y_{t-1} + a_{\text{lead}} E_t y_{t+1} - a_r (i_t - E_t \pi_{t+4}^4 - \bar{r}_t) + \varepsilon_t^y, \quad (2)$$

with $\bar{r}_t$ the neutral real rate of Section 2; the Phillips block is

$$\pi_t = \lambda_1 E_t \pi_{t+4}^4 + (1 - \lambda_1) \pi_{t-1}^4 + \kappa y_{t-1} + \varepsilon_t^\pi, \quad (3)$$

so that λ_1 is the weight on the four-quarter-ahead expectation and $1 - \lambda_1$ the weight on lagged four-quarter inflation, and the block is purely forward-looking at $\lambda_1 = 1$; the policy slot carries the IFB rule

$$i_t = \max \left\{ \underline{i}, \rho i_{t-1} + (1 - \rho) [\bar{r}_t + \pi_{t+3}^4 + c_\pi (\pi_{t+3}^4 - \pi^*) + c_y y_t + c_0] + u_t \right\}, \quad u_t = \phi_u u_{t-1} + \eta_t, \quad (4)$$

with smoothing ρ , responses c_π and c_y to the four-quarter inflation forecast three quarters ahead and to the contemporaneous gap, an intercept c_0 , a persistent shock u_t with autocorrelation ϕ_u and innovation standard deviation σ_η , and the floor $\underline{i}$ at 0.125 percent, the model's ELB. Expectations are model-consistent, as in the book, whose agents hold model-consistent expectations and no other kind; the anchored, cognitive-discounting, and learning families of Section 5 are this paper's, each nesting the book's case. At each quarter the perfect-foresight equilibrium over a forward horizon of 48 quarters is solved with the floor as a complementarity condition, the structural shocks at zero over the horizon except where a section feeds them, the gap at zero and inflation at the target beyond it, and the neutral-rate path over the horizon read from the path of Section 2 from that quarter forward and held at its final value, 1.01 percent, beyond the record; the agent's rule of Section 7 takes the same form with u_t set to zero. The expectations block admits a family of departures, stated in Section 5, that nest the model-consistent block at one parameter value.

4.2 Estimator and certification

The parameters are estimated by a likelihood filter over the observation rows of Section 2, the staff gap rows, the staff inflation expectation, realized inflation, the funds rate, the two-year rate, and the four-quarter survey, with the equilibrium re-solved at every quarter and a penalty for any quarter at which no equilibrium exists at the trial parameters. Standard errors come from the finite-difference Jacobian of the standardized innovations. Before the estimator was applied to the record it was certified on a synthetic oracle: 18 conditions, the final four of which carry a scripted floor spell and

¹² Ali Alich, Kevin Clinton, Charles Freedman, Michel Juillard, Ondra Kamenik, Douglas Laxton, Jarkko Turunen, and Hou Wang, "Avoiding Dark Corners: A Robust Monetary Policy Framework for the United States," IMF Working Paper 15/134 (Washington, DC: International Monetary Fund, 2015), to which the chapter's own footnote refers for the model's equations. Its output-gap equation responds to a long real rate, a weighted average of the current short real rate and its expected averages over the next four, twelve, and twenty quarters, carries a zero coefficient on the current short real rate, and carries a bank-lending-tightening term. The IS block here responds to the current short real rate alone and carries no credit term; both are reductions made for estimation on the real record, and the coefficients reported for the IS block in the book-economy column of Table 1 are this paper's calibration of that reduced block, not values the chapter prints.

the final three among them the record's missing-row pattern as well, and 75 estimation units on economies simulated at known parameters. Under a floor spell with every row present (condition O15) the anchored estimator recovers the anchoring weight at 0.769 to 0.809 on five seeds against a truth of 0.80, with standard errors of 0.037 to 0.046, and with the record's missing-row pattern as well (O16) at 0.737 to 0.805 with standard errors of 0.039 to 0.048; the floor spell supplies the identification the pre-floor window lacked. Appendix A lists the conditions.

4.3 The multi-start search and the estimate of record

The likelihood of the block of record has more than one basin, and along one direction it has no floor the record supplies. The direction is the scale of the policy block. With the rule's inflation response c_π held at each point of a grid registered before the profile was computed, 1, 2, 3, 4, 6, 8, 12, and 16, and every other free parameter re-optimized from two catalogued starts, the objective per observation row is 0.1899, 0.1549, 0.1445, 0.1511, 0.1416, 0.1349, 0.1450, and 0.1543: a local minimum at 3, a lower one at 8, about 8 units of log-likelihood deeper, and a rise beyond it only because the intercept reaches its bound from 8 on. Along the way the intercept and the policy shock's standard deviation move with c_π (at 8 they stand at 6.94 and 1.60), which is the signature of a ridge: the moments fix the policy block's fit to the funds rate and not its scale. The estimate obtained before the profile had sat at the upper bound of its box, c_π of 8.0. Neither reporting a bound as an estimate nor widening the box after seeing where the estimate went meets the standard of a registered search, and the scale is fixed instead by a convention registered before any estimate under it was computed: $c_\pi = 3.0$, the point at which the earlier search's deep basins sat, at which the profile has its interior minimum, and which the profile's registration named as its interior point before the profile's shape was known; the book's value of 0.91, whose neighborhood the profile rejects, is not a candidate. Every estimate of the world model on the record in what follows, the full window's, the frozen window's, and each rolling window's, holds c_π at 3.0 and carries no standard error for it. What the agent learns does not turn on the choice: an agent trained on the family re-polished at 3, before the convention, prescribed rates above the FOMC's over 2016 on (0.57 points against 0.97 at the bound) and learned an inflation response of 1.48 against 1.77; the size of the paper's distance from the FOMC does turn on it, and Section 9 states it as a limitation.

Under the convention the search ran from 40 random starts drawn uniformly over the estimator's box on a catalogued stream and from four warm starts registered from the records of earlier constructions: the profile's point at 3, the deepest basin of the search before the convention, and the estimates of record of two earlier constructions, each with c_π at the convention's value. The deepest basin, at 0.1445 per row, was reached by two warm starts, the profile's point and the earlier search's deepest basin, and by none of the random starts. A second family, 20 starts (18 random and two warm) at 0.1583 to 0.1612, lies about 12 units of log-likelihood above it with a Phillips block of the same kind, λ_1 from 0.91 to 0.95 and a slope from 0.053 to 0.062; the remaining 22 random starts stop at 0.1993 to 0.3169, with λ_1 from 0.62 to 0.95 and, in nine of them, a Phillips slope at zero. The estimate of record is the deepest basin re-polished under a box widened in λ_1 and c_0 , from the warm start W2. No parameter lies at a bound of its box, and λ_1 stands at 0.946, short of its structural limit; the estimate lies on one joint constraint, the two weights of the IS block lie on the cap $a_{\text{lag}} + a_{\text{lead}} \leq 0.97$ that the estimator imposes for a stationary solution, as they do in the earlier estimate of Section 2, the frozen estimate, and each rolling window's. Their standard errors are computed at that point as if the two weights were free, so that they differ although the constraint makes the weights' moves equal and opposite. Of the multi-start search's five registered predictions, the two that bear on identification divided: the deepest basin was reached from at least

two starts, as predicted, and the random starts alone did not reach it, against the prediction that they would, the deepest of the 40 stopping at 0.1583. The estimate of record was found through the record's own catalogued points, not by the random search, and the paper states that as the second half of the identification evidence beside the profile. Figure 3 shows the 44 optima, and Table 1 reports the estimate of record beside the best estimate from the data-driven grid start, run B, and the book calibration.

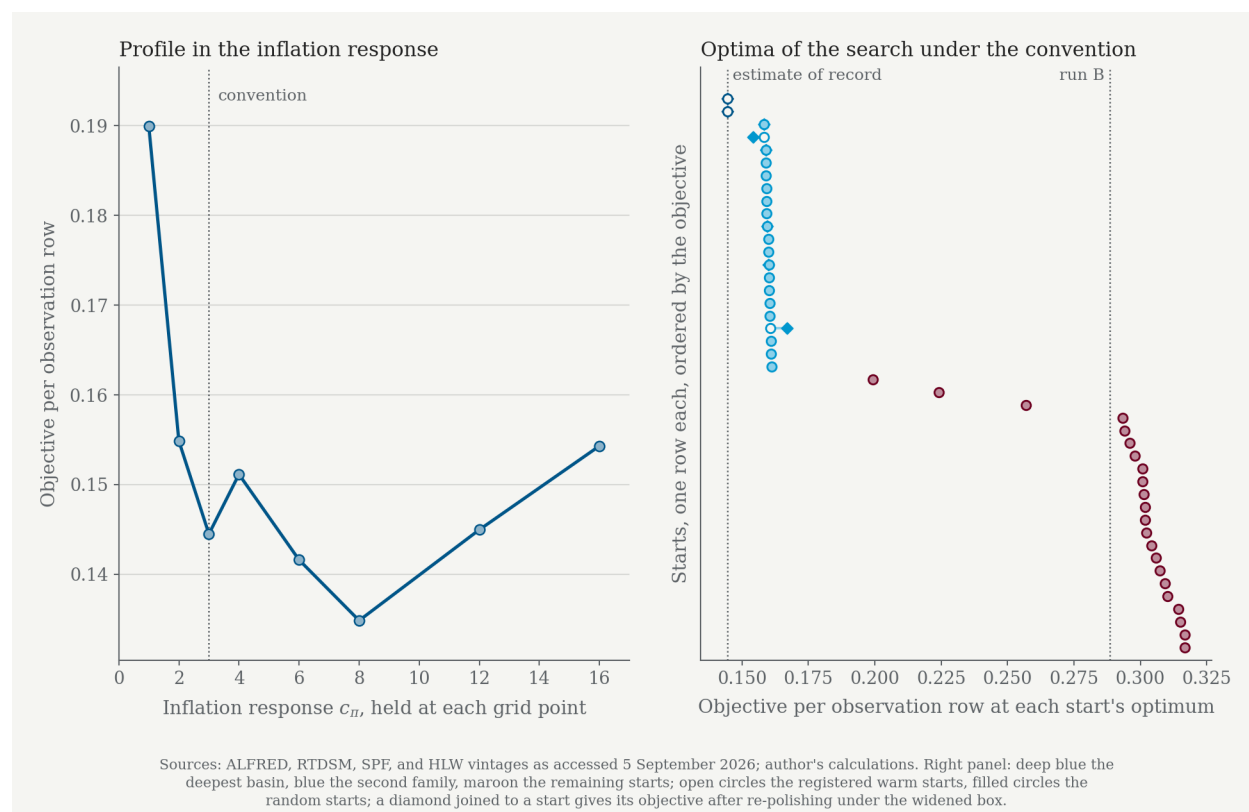


Figure 3: The profile in the inflation response and the optima of the search, 1996:Q1–2026:Q2. Left panel: the objective per observation row (lower is better) with c_π held at each point of the registered grid and every other free parameter re-optimized, the better of two catalogued starts at each point; the dotted line marks the convention's value, 3.0. Right panel: the 44 starts of the search under the convention, ordered by the objective at each start's optimum under the estimator's box, one row each, the vertical position the start's rank and not a value; deep blue marks the deepest basin, blue the second family, maroon the remaining starts, with open markers for the four registered warm starts and filled markers for the 40 random starts; for the starts that were re-polished under the widened box a second marker, joined to the first, gives the re-polished objective. Dotted lines mark run B (0.2885) and the estimate of record (0.1445). Sources: ALFRED, RTDSM, SPF, and HLW vintages as accessed 5 September 2026; author's calculations.

Table 1: The estimate of record of the world model, 1996:Q1–2026:Q2. Run B is the best estimate reached from the data-driven grid start; the estimate of record is the deepest basin of the search of Section 4.3 re-polished under the widened box, reached from a registered warm start; in both the inflation response c_π is held at the convention’s value of 3.0 and carries no standard error; SE is the standard error of the estimate of record from the finite-difference Jacobian; the book-economy column is the calibration used for the book-economy experiments of this paper: λ_1 and κ as chapter 12 prints them, the IS coefficients this paper’s, the chapter printing no IS equation (Section 4.1). The real-rate coefficient is calibrated. Sources: ALFRED, RTDSM, SPF, and HLW vintages as accessed 5 September 2026; author’s calculations.

Parameter	Symbol	Run B	Estimate of record	SE	Book economy
IS: lagged gap	a_{lag}	0.684	0.771	0.030	0.900
IS: expected gap	a_{lead}	0.286	0.199	0.034	0.050
IS: real rate (calibrated)	a_r	0.100	0.100	—	0.100
Phillips: expectation weight	λ_1	0.662	0.946	0.041	0.700
Phillips: slope	κ	0.0314	0.0596	0.0166	0.1000
Rule: smoothing	ρ	0.839	0.656	0.027	—
Rule: inflation response	c_π	3.00	3.00	—	—
Rule: gap response	c_y	1.58	0.84	0.28	—
Rule: intercept	c_0	0.48	3.00	0.80	—
Rule shock persistence	ϕ_u	0.790	0.959	0.016	—
Rule shock sd	σ_η	0.496	0.947	0.103	—
Survey noise sd	σ_{spf}	0.473	0.798	0.077	—

The estimate of record and run B are two accounts of the funds rate and of inflation. Run B, the best estimate the data-driven grid start reaches, carries a Phillips block that leans on lagged inflation (λ_1 of 0.662) with a flat slope (0.0314), and a heavily smoothed rule with a moderate policy shock; the estimate of record carries a Phillips block nearly forward-looking (λ_1 of 0.946) with a slope of 0.0596 (standard error 0.0166), a gap response of 0.84, and a large and persistent policy shock (ϕ_u of 0.959, σ_η of 0.947). Its objective lies below run B’s by about 120 units of log-likelihood. The equilibrium exists at every quarter of the record under the estimate, both floor spells included, and the standardized innovations over the floor quarters have 1.28 times the standard deviation of those elsewhere.

4.4 The model’s forecasts against the staff’s

A world model that forecasts worse than the forecaster it is compared with would be a weak place to train a policy, so Table 2 scores the estimate of record against the staff at the 84 decision points from 2000:Q1 to 2020:Q4 that carry an in-quarter Tealbook vintage, 2020:Q4 being the final decision point at which the record admits one. At each, the model’s one-quarter-ahead gap and inflation and its four-quarter-ahead inflation are computed from the filtered state under the estimated policy block, on the blocks alone and with the residual components; the staff’s projections are read from the same vintage; and the realizations are read from the panel’s vintage five quarters on, the staff’s own real-time gap for the gap and the core PCE tail for inflation. The staff forecasts better on every line: one-quarter inflation at a root-mean-square error of 0.640 against the model’s 0.772 on the blocks and 0.731 with the components, four-quarter inflation at 0.562 against 0.745, the gap at

1.218 against 1.345 and 1.301. The margins are those of a forecaster who sees within-quarter data the model's state does not: the four-quarter ratio is 1.32, the loss differentials on inflation are within two of their standard errors of zero at one quarter with the components and just outside at four, and the gap differential is not resolved. The components move every model line toward the staff's. Of the three registered predictions, the two that the staff would lead at one quarter, on inflation and on the gap, held; the third, that the model's four-quarter inflation error would lie within a quarter of the staff's, failed, at a ratio of 1.32.

Table 2: The model of record against the staff at the 84 decision points from 2000:Q1 to 2020:Q4 with an in-quarter Tealbook vintage. One ahead and four ahead are the horizons in quarters; RMSE and mean error are of the forecast against the realization, in percentage points (the gap in percent of potential); the loss differential is the mean of the model's squared error less the staff's, with its standard error over the quarters. Realizations are read from the panel's vintage five quarters after the decision point: the staff's real-time gap, and quarterly core PCE inflation from that vintage's complete quarters; the four-quarter line has 83 quarters. Sources: Philadelphia Fed Tealbook and RTDSM vintages and ALFRED vintages as accessed 5 September 2026; author's calculations.

Forecast	Quarters	Model		Staff		Loss differential
		RMSE	Mean err.	RMSE	Mean err.	Mean (SE)
Gap, one ahead, blocks	84	1.345	+0.486	1.218	-0.008	+0.325 (0.440)
Gap, one ahead, components	84	1.301	+0.250	1.218	-0.008	+0.209 (0.404)
Inflation, one ahead, blocks	84	0.772	+0.071	0.640	+0.061	+0.186 (0.076)
Inflation, one ahead, components	84	0.731	+0.045	0.640	+0.061	+0.124 (0.067)
Four-quarter inflation, four ahead	83	0.745	+0.100	0.562	-0.013	+0.238 (0.108)

5 The Expectations Block Against Three Families

Model-consistent expectations are the construction under which forward guidance at a floor has its largest effects, and the one the behavioral and learning literatures have modified for that reason. Three modifications were tested, each nesting the block of record, each certified on the oracle under a floor spell, and each estimated on the full window with the likelihood profiled in the family's parameter. Table 3 collects the estimates.

Table 3: The expectations families on the full window. The parameter is the family's departure from model-consistent expectations, at 1 in every case; the penalty is the increase in the objective, in units of log-likelihood over 836 rows, when the parameter is held at 0.8 with every other free parameter re-optimized. The learned-anchor family's weight is on the model-consistent expectation against the learned level. Sources: as Table 1; author's calculations.

Family	Parameter	Estimate	SE	Penalty at 0.8
Anchored, level at the target	w	1.000	0.042	39.8
Anchored, level as measured	w	1.000	0.044	—
Cognitive discounting	m	1.000	0.018	110.9
Learned anchor	w	1.000	0.008	—

The anchored family places the four-quarter expectation at a weight w on the model-consistent expectation and $1 - w$ on a level, first the target and then the level measured from the ten-year survey (1.91 in core PCE units, standard deviation 0.21 over 122 surveys). On the pre-floor window the likelihood was indifferent to w ; on the full window w sits at its bound with a standard error of 0.044, and holding it at 0.9 and 0.8 costs 32.0 and 39.8 units of log-likelihood. The cognitive-discounting family attenuates every forward expectation k quarters ahead by m^k toward its default; the oracle recovers a discount of 0.85 at 0.828 to 0.931, and on the record m sits at one, with the profile 50.8, 80.2, and 110.9 units worse at 0.95, 0.9, and 0.8. The fit at the floor hardly moves under discounting: the floor quarters' innovation ratio is 1.60 at one and 1.58, 1.59, and 1.66 at 0.95, 0.9, and 0.8, while the fit of the whole record worsens by the profile's units. The learned-anchor family makes the level a constant-gain mean of realized inflation observed by the ten-year survey; the oracle recovers a gain of 0.05 at 0.0483 to 0.0534 with standard errors below 0.0028. On the record the gain is 0.030 (0.0024), a half-life near 23 quarters, and the learned level runs from 2.85 to 2.61 over the window, so the long-run survey behaves as a learner's anchor; but the weight on the model-consistent expectation is 1.000 (0.008), and no lower than 0.996 at any gain the profile holds. Given a freely moving anchor to lean on, the agents in the macro rows lean on none of it. Figure 4 plots the three profiles, and Figure 5 plots the survey against the learned anchor and realized inflation: the survey's level runs below the learner's mean over the first years of the window and holds nearer the target than the learner does after 2021, so the survey is more anchored than a constant-gain mean of experience, and it is still the survey that learns rather than the agents.

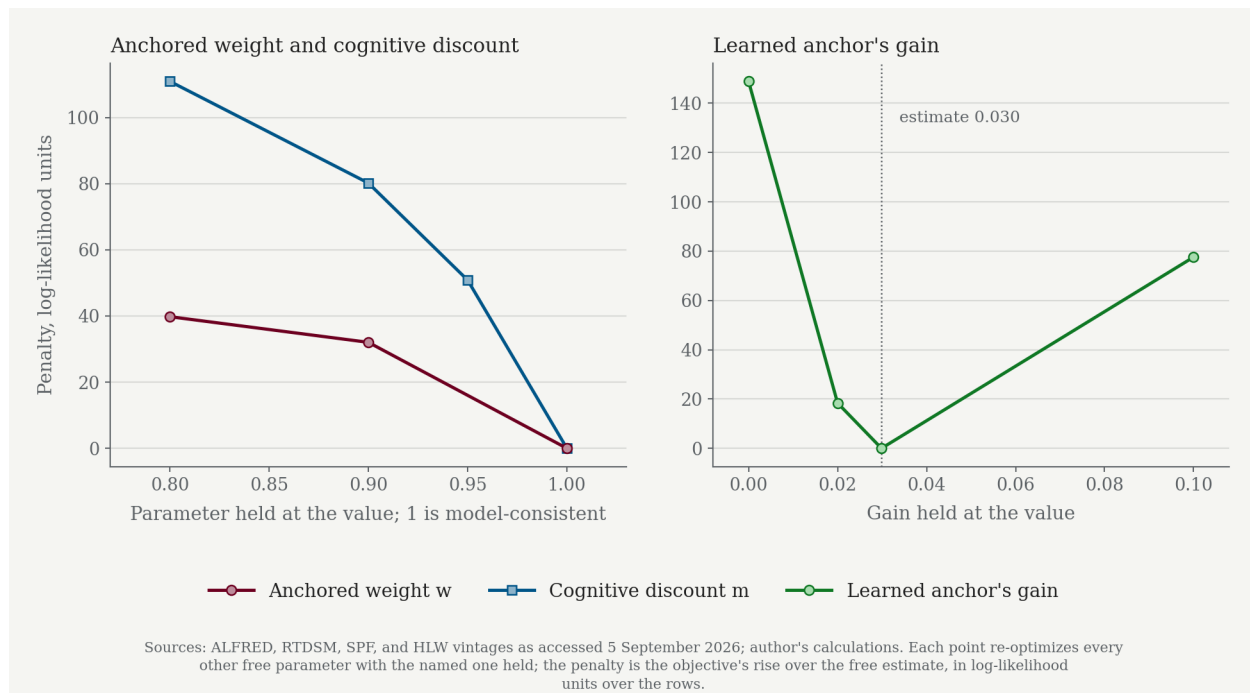


Figure 4: The likelihood profiles of the three expectations families on 1996:Q1–2026:Q2. Left panel: the penalty in log-likelihood units over 836 rows when the anchored family's weight w (maroon circles) or the cognitive-discounting family's discount m (deep blue squares) is held below one with every other free parameter re-optimized; both are zero at one, where each family is model-consistent, and rise to 39.8 and 110.9 at 0.8. Right panel: the penalty when the learned anchor's gain (green circles) is held at zero, 0.02, and 0.10 against its estimate of 0.030 (dotted line): 149 at zero and 78 at 0.10. Sources: as Table 1; author's calculations.

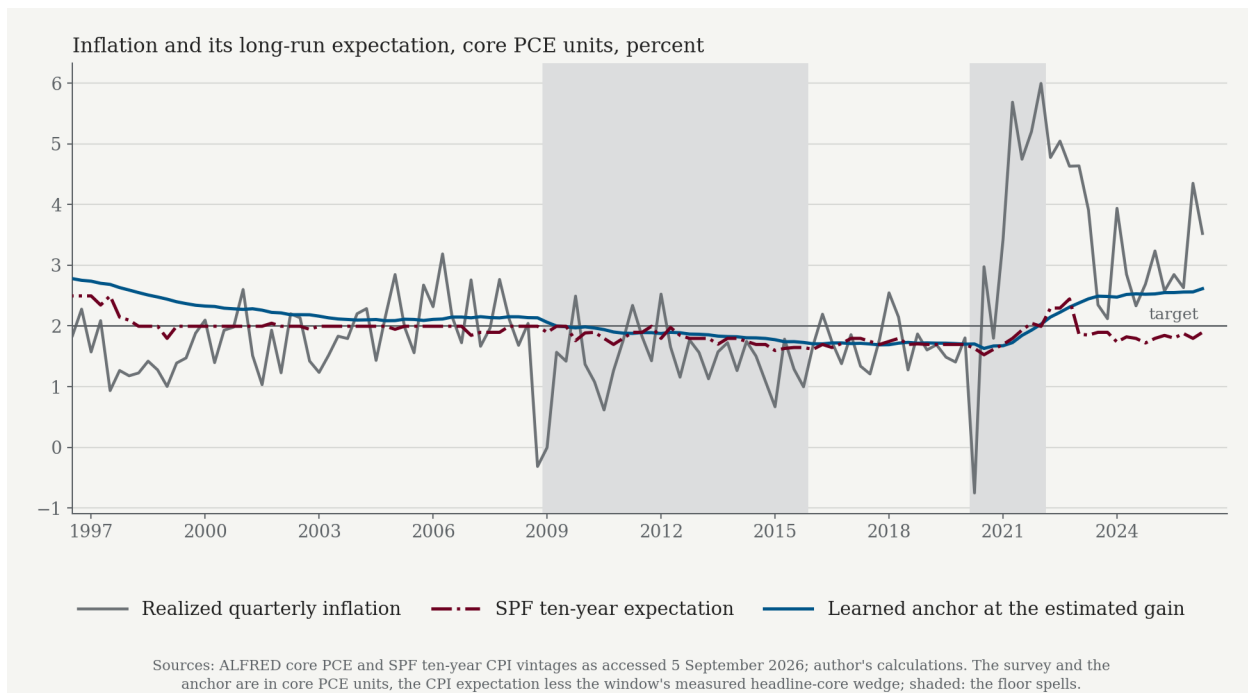


Figure 5: The long-run survey expectation and the learned anchor, 1996:Q1–2026:Q2. The maroon dash-dot line is the SPF ten-year CPI inflation expectation in core PCE units, the survey less the window's measured headline-core wedge; the deep blue line is the learned anchor, a constant-gain mean of realized quarterly inflation at the estimated gain of 0.030, starting from 2.85 and ending at 2.61; the gray line is realized quarterly core PCE inflation at an annual rate; the thin gray rule the target; shaded spells the quarters at the floor. Sources: ALFRED core PCE and SPF ten-year CPI vintages as accessed 5 September 2026; author's calculations.

6 Residual Components and a Floor-Aware Null

The blocks leave residuals, and a learned correction can fit them: a GP component on the filtered states for each of the IS and Phillips shocks, with an isotropic kernel, log-normal priors, maximum a posteriori (MAP) hyperparameters, and the signal standard deviation capped at a detection floor calibrated on the oracle. The certification statistic for such a component is the leave-one-out (LOO) log-score excess of the GP over a Gaussian with no structure. Its null distribution was obtained on 25 zero-truth trajectories simulated with the floor spell and the record's missing-row pattern, and it is not the pre-floor null: the median excess for the Phillips component is 0.217 and its 95th percentile 0.385, against a pre-floor 95th percentile of 0.079. A floor spell puts structure into the filtered shocks at the true parameters, because the rate row is censored there and the floor marks those rows in the state, and a GP on the state features reads it as signal.

On the estimate of record the Phillips component's excess sits at the 48th percentile of that null, carrying no structure beyond what a floor spell induces; the IS component's excess, 0.239, lies above the null's 95th percentile of 0.145. The estimate of record accounts for the rate and inflation rows by a rule with a persistent shock and leaves structure in the demand residual, which the paper states as a limitation of the IS block. The components are not part of the agent's training environment, for the reason Section 7 establishes; where they are used, in the existence tests of

the policy rollouts, 36 fitted rows sit at the floor and 0 of 24 rollout quarters lie outside the fitted support, so the equilibrium exists on every policy the rollouts evaluate.

7 The Policy Agent

The agent is a policy rule trained by gradient descent inside the world model; this section states its loss and training environment, decomposes what an agent trained with the residual components exploits, and selects the rule of record by a search over training streams.

7.1 Loss and training environment

The agent minimizes the loss of Adrian, Laxton, and Obstfeld over a 48-quarter horizon: the sum of squared four-quarter inflation deviations from target with weight α of one, squared output gaps with weight β of one, and squared quarterly rate changes with weight γ of one half. Its policy is an IFB rule, four coefficients (ρ, c_π, c_y, c_0) trained by gradient descent on episodes that start from the real-time states of the record, 118 starts from 1997:Q1 through 2026:Q2, 36 of them at the floor, with the equilibrium solved exactly along each episode and the floor as a complementarity condition.

7.2 Structured or learned?

What part of a trained policy's advantage comes from the structure it was trained against, and what part from the learned correction? An agent trained with the components in the environment, sampling from their posterior on the blocks of run B, sits nearer the FOMC than the agent of record: its mean difference from the realized rate over the record is +0.55 with an RMS difference of 1.00, against the agent of record's +0.768 and 1.026, and before the first floor spell its RMS difference is 0.53 against 0.82; it holds the floor on 27 of the 36 floor quarters. Its advantage over the book rule depends on where it is measured. In its own blocks' exact system, without the components, its realized loss from the book state is 9.08 against the book rule's 7.13, a loss; in the book calibration it wins, 13.33 against 14.38; inside the model with the components its loss is 35.8 against the book rule's 124.5. Figure 6 measures the advantage across the estimate's uncertainty rather than at one point: over 100 draws of the estimated blocks from the estimate of record's asymptotic distribution, the agent of record beats the book rule on 100 draws, by 4.24 on average, and the components-trained agent on 43, by 0.41 on average, the first profile above the second at every threshold; the three registered predictions held. An advantage that lives in the components, which Section 6 found carry no signal in the inflation block, does not survive the estimate's uncertainty on more than half the draws; the training environment of record is the blocks alone, and the policy nearest the FOMC's is the one trained on a correction the record does not support as signal.

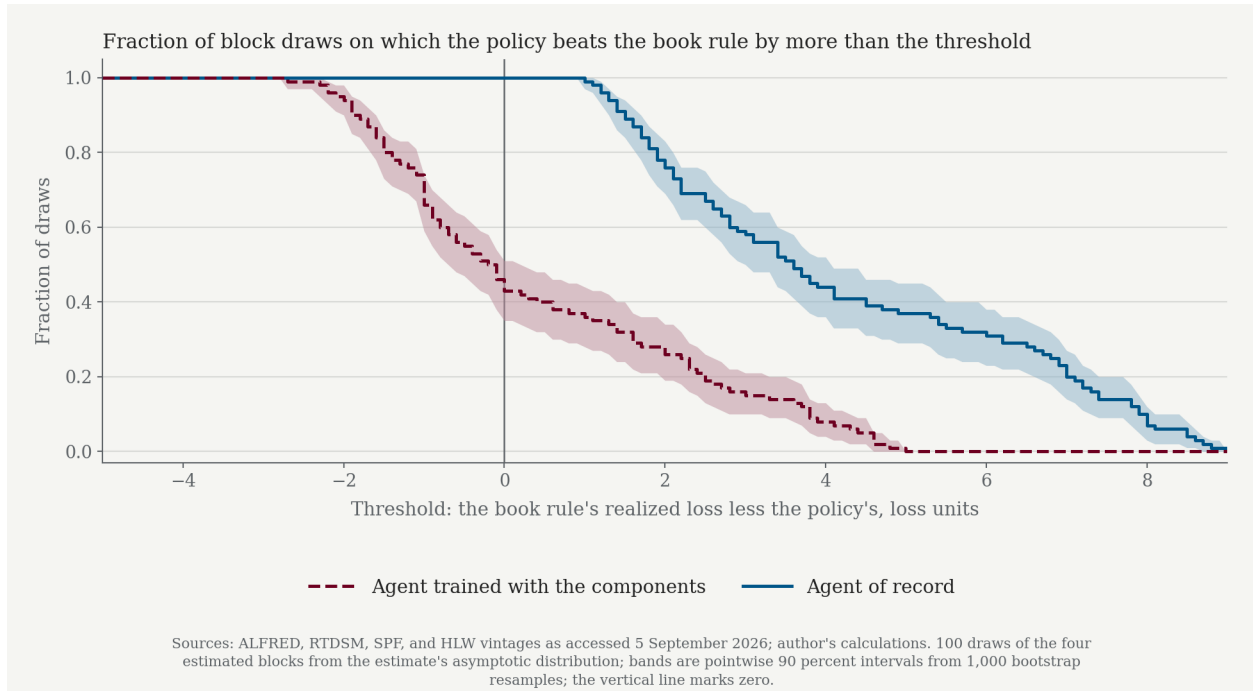


Figure 6: The performance profile of the exploitation bound. For each of 100 draws of the four estimated blocks from the estimate of record's asymptotic distribution, the exact economy at the draw gives the realized loss from the book state of the agent of record (deep blue, solid), the agent trained with the residual components (maroon, dashed), and the book rule; the curve is the fraction of draws on which the policy beats the book rule by more than the threshold on the horizontal axis, and the shaded band the pointwise 90 percent interval from 1,000 bootstrap resamples; the vertical line marks zero. Sources: as Table 1; author's calculations.

7.3 Training streams and the checkpoint of record

A training run of a fixed number of steps leaves the coefficients set by the budget, and a single run leaves them set by its stream: the draws of blocks and starts a run sees are one sample, and another sample ends elsewhere. The agent of record is the best checkpoint of a registered search. Each training episode draws the four estimated blocks from the estimate's asymptotic distribution, so the policy optimizes the expected loss over the estimation uncertainty rather than the loss at the point estimate. The search trains eight streams of draws, each catalogued before training, for a registered horizon of 2,000 steps each; a validation set of 32 held-out block draws and eight held-out starts, fixed before training, is read every 10 steps on every stream; the checkpoint of record is the lowest validation reading over the streams and their steps, and the horizon is extended only if that reading falls in its final 200 steps. The validation starts are drawn uniformly from the pool of starts on a catalogued stream of their own, for the agent of record 2000:Q1, 2010:Q4, 2011:Q2, 2013:Q3, 2016:Q1, 2016:Q4, 2020:Q4, and 2022:Q2, seven of them from 2008:Q4 on; eight test starts and 32 test draws are drawn from the rest and read by no selection.

On the estimate of record the validation loss fell from 53.47 at the initial coefficients to 38.55 at the checkpoint of record, stream 8 at step 450, and the horizon was not extended (Figure 7). On the test set the selected stream ranks third of eight, its test loss 0.12 above the best stream's.

The streams' validation ranking says little about their test ranking: over the agent of record's eight streams the Spearman correlation of the checkpoints' validation and test losses is 0.02, and over the eleven searches of the paper, the agent of record's, the four of Table 10, the four of Table 9, run B's, and the frozen agent's, it is positive in two, from -0.31 to 0.64 across the four members, -0.67 to -0.14 across the weight variants, -0.93 for run B's agent, and -0.90 for the frozen agent. The streams' checkpoints spread by 0.87 on the validation set and by 1.40 on the test set, on a validation loss of 38.55 at the checkpoint of record, which the few quarters far from target dominate; the checkpoint of record is the best of eight on its own validation set and on nothing else, and Section 8.2 reads on the 82 rolling searches what a ranking of eight streams on a set of this size carries to a fresh one: the test ranking carries, and the validation ranking does not. The counts are not independent readings: nine of the eleven searches, the agent of record's, the members', and the weight variants', share one estimate, one validation set, and one test set, and run B's correlation rests on test losses within 0.2 percent of one another. The rules at the eight checkpoints differ by more than their losses: across the agent of record's streams c_π runs from 1.79 to 2.37, c_y from 1.37 to 2.38, and ρ from 0.64 to 0.69, while their test losses lie within 8.7 percent of one another. The loss changes little along a ridge of rules, the selection names one point on it, and the coefficients of record are read as that point, not as the only rule the construction supports. A criterion that pools a checkpoint's validation and test losses, registered after these test losses were seen and read against a fresh held-out set, selects the same stream for the agent of record and another in three of the eleven searches, the frozen agent's and two of the weight variants', each of them ranking higher on the fresh set than the stream the construction selected; the rule of record is the construction's selection throughout. The search on run B's blocks was extended once, to 4,000 steps, its checkpoint at step 3,470 of stream 5.

The validation set replaced an earlier one. The earlier rule kept the eight earliest distinct quarters of thirty-two draws, so the agent of record's validation starts ran from 1997:Q1 to 2006:Q4, before either floor spell, while the test starts spanned the record. On the agent of record the two rankings then agreed (Spearman 0.93, on validation losses near 4.19), but on the frozen agent they ran opposite (-0.90) and across the 82 rolling windows the median was -0.26 , with the selected stream seventh or eighth on the test set in 44 of them. The replacement was registered before any run under it, with six predictions. It moved the agent of record's loss in the exact economy at the estimate of record by 0.0376, more than the 0.0180 spread of the five searches under the earlier set, against the prediction; the agent of record's selected stream stays in the better half of the test ranking, as predicted, and the frozen agent's stays last, against the prediction; the three predictions on the rolling windows are reported in Section 8.2. The paper reports the checkpoint the construction selects with its test rank beside it.

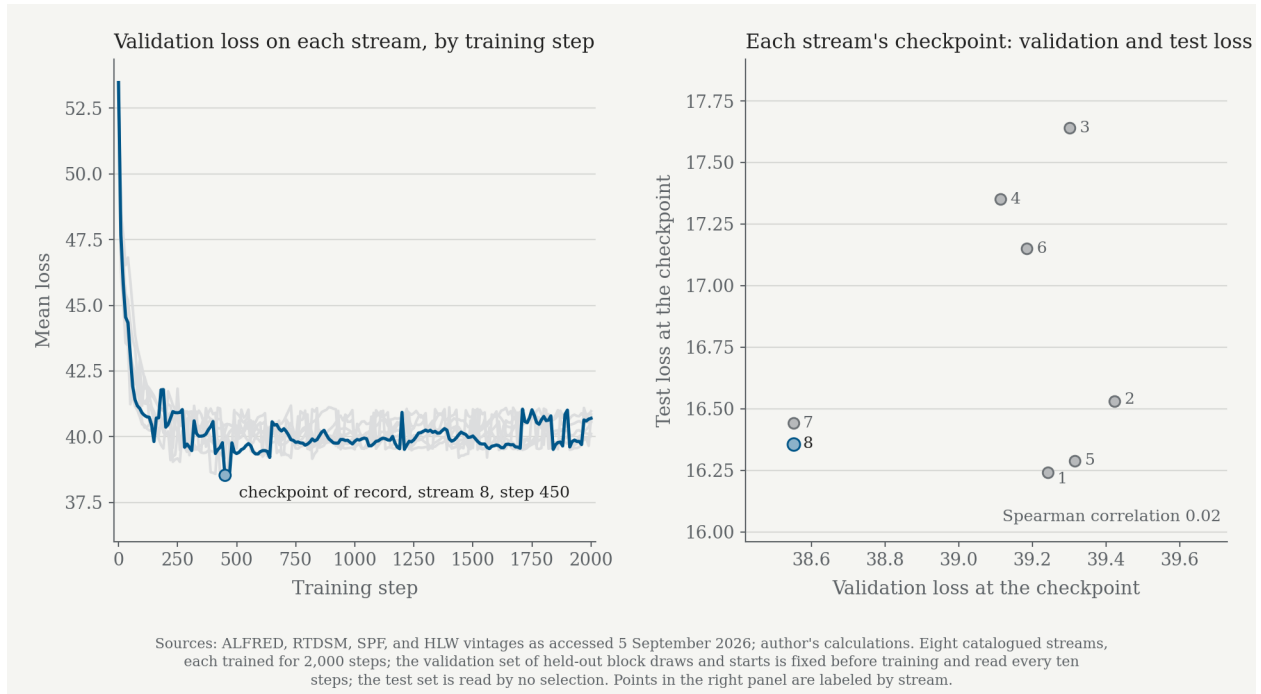


Figure 7: The search over training streams on the estimate of record. Left panel: the mean loss over the fixed validation set, 32 held-out draws of the estimated blocks times eight held-out real-time starts, read every 10 steps on each of the eight streams over the horizon of 2,000 steps; the selected stream, 8, is deep blue and the others light gray, and the marker is the checkpoint of record at step 450. Right panel: each stream's checkpoint, its validation loss against its loss on the held-out test set, the checkpoint of record in deep blue; the Spearman correlation of the two is 0.02. Sources: as Figure 3; author's calculations.

The agent of record is ρ of 0.644, c_π of 2.165, c_y of 2.176, and c_0 of -0.246 . It beats the book rule in its own exact world, 5.92 against 9.94, and in the book calibration, 10.34 against 14.38; training under the estimate's uncertainty transfers to the calibration better than training at the point estimate did (11.01).

8 The Optimal Policy and the FOMC

This section sets the agent of record against the FOMC's realized decisions at every real-time state, with bands from the estimate's uncertainty and beside three published rules; repeats the comparison with everything frozen at the end of the first floor spell; and simulates the outcomes each rule would have produced with the record's own shocks.

8.1 The agent of record against the FOMC

At each of the 118 real-time states the agent of record's first-quarter policy rate is computed in the exact economy at the estimate of record, with a band from the 10th to the 90th percentile over 20 draws of the estimated blocks from the estimate's asymptotic distribution, and compared with the realized funds rate at that decision point and with the book rule's rate. At three of the 2,360 pairs of

a state and a draw, at 2009:Q4 and 2010:Q2, no equilibrium exists under the draw, and the band at those quarters is read over the draws that have one; the existence of the equilibrium under a draw at the first floor spell sits at the solver's convergence edge, and the count is the one computed on the author's machine, where the record was built. Figure 8 plots the three rates and the band, and Table 4 reports the comparison by period.

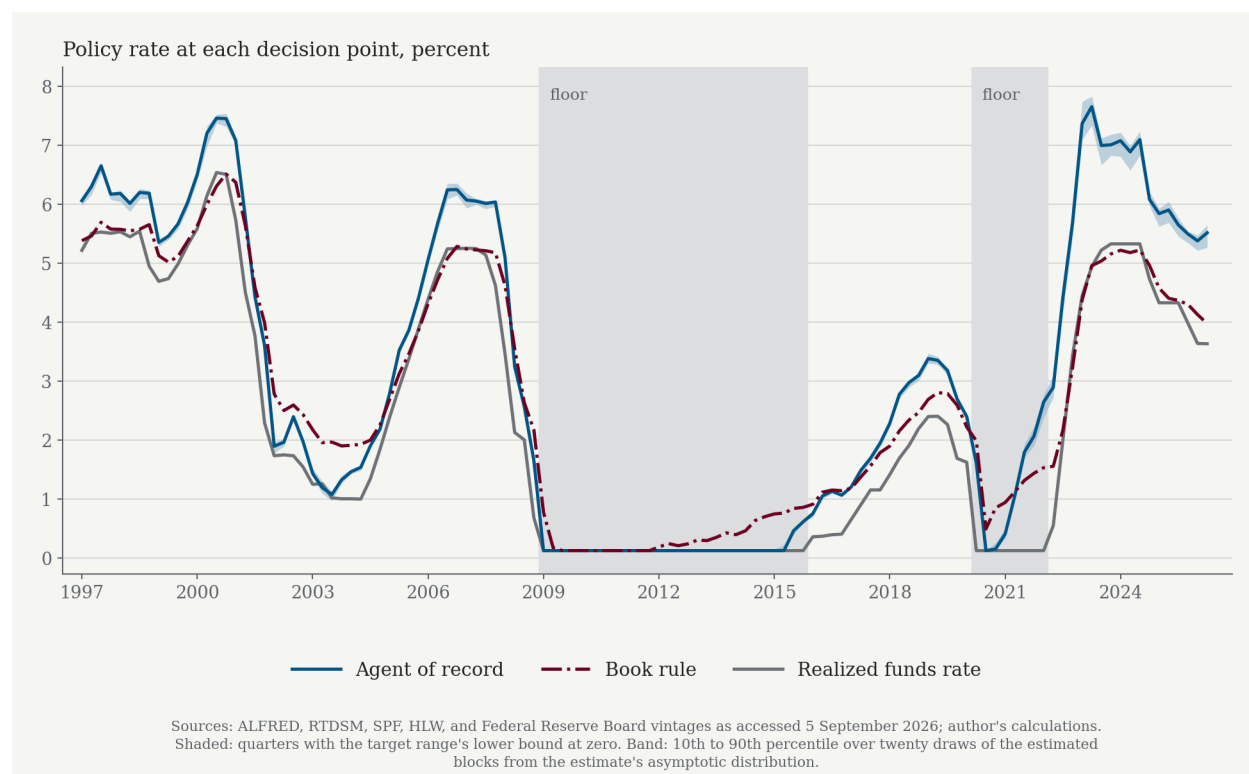


Figure 8: The agent of record against the FOMC at the 118 real-time decision points from 1997:Q1 to 2026:Q2. The deep blue line is the agent's first-quarter policy rate in the exact economy at the estimate of record and the blue band its 10th to 90th percentile over 20 draws of the estimated blocks; the maroon dash-dot line is the book rule's rate; the gray line is the realized funds rate at the decision point, at the model's floor where the target range's lower bound was zero (the shaded spells). Sources: ALFRED, RTDSM, SPF, HLW, and Federal Reserve Board vintages as accessed 5 September 2026; author's calculations.

Table 4: The agent of record against the FOMC, 1997:Q1–2026:Q2. Mean is the mean difference between the policy’s first-quarter rate and the realized funds rate at the decision point, in percentage points; RMS is the root mean square of that difference; coverage is the share of quarters at which the realized rate lies within the agent’s 10th–90th percentile band over 20 block draws; below is the number of quarters at which the realized rate lies below the band. The book rule is the IFB rule at the book coefficients. Sources: as Table 1; author’s calculations.

Period	Quarters	Agent		Book rule		Coverage	Below
		Mean	RMS	Mean	RMS		
All quarters	118	+0.77	1.026	+0.40	0.602	0.25	88
Pre-floor, 1997:Q1–2008:Q4	48	+0.74	0.820	+0.45	0.665	0.04	46
Floor, 2009:Q1–2015:Q4	28	+0.03	0.113	+0.23	0.343	0.93	2
Floor, 2020:Q2–2022:Q1	8	+1.11	1.418	+1.09	1.169	0.25	6
From 2016:Q1, second floor spell excluded	34	+1.34	1.490	+0.33	0.461	0.00	34
2021:Q1–2023:Q4	12	+1.96	2.085	+0.52	0.813	0.00	12

The mean difference over all 118 quarters is +0.768 with a standard error of 0.063 under independence; the quarterly differences are serially correlated, and a Newey–West standard error with 4 lags is 0.128, with a moving-block bootstrap of 2,000 resamples in blocks of 8 quarters giving a 95 percent interval of +0.451 to +1.065. The band covers the realized rate on 0.25 of quarters and the realized rate lies below it on 88. The one exception is the first floor spell: on 2009:Q1–2015:Q4 the agent’s band has its floor at the floor on 26 of 28 quarters and its RMS difference is 0.113 against the book rule’s 0.343, and its mean difference of +0.030 carries a Newey–West standard error of 0.027, the closest agreement with the FOMC anywhere in the record. On the second spell the agent sits +1.11 above the floor on average, and over 2021:Q1–2023:Q4 it sits +1.96 above the realized rate, below the band on every quarter. These are properties of the optimal policy under the loss of record in the estimated world, conditional on the loss weights and on the world model, and the paper states them as such.

Table 5 places the agent beside the three rules the FOMC reports in its Monetary Policy Report, each evaluated at the same real-time states in output-gap form on the state’s four-quarter core inflation, gap nowcast, and lagged rate, bounded below by the model’s floor: the Taylor (1993) rule, the balanced-approach rule with a unit response to the gap, and the first-difference rule. The agent sits between the FOMC and the level rules. On all quarters the Taylor (1993) rule prescribes +1.48 above the realized rate and the balanced-approach rule +1.60, against the agent’s +0.768; over 2021:Q1–2023:Q4 the two level rules prescribe +6.12 and +7.24 above the realized rate, above it on all 12 quarters, where the agent prescribes +1.96. The agent’s rate correlates with the first-difference rule’s at 0.98 and with the book rule’s at 0.97, against 0.72 with the Taylor (1993) rule: the optimal rule under the loss of record is inertial, closer in form to the first-difference rule than to a level rule, and its distance from the FOMC is a fraction of the level rules’ distance.

Table 5: The agent, the book rule, and three published rules against the FOMC, 1997:Q1–2026:Q2. Mean and RMS are the mean and root-mean-square difference between the rule’s rate and the realized funds rate at the decision point, in percentage points, over the quarters of Table 4. Taylor (1993) is $i = 2 + \pi + 0.5(\pi - 2) + 0.5y$ on four-quarter core inflation π and the gap nowcast y ; the balanced-approach rule has a unit gap response; the first-difference rule adds $0.5(\pi - 2) + 0.5(y - y_{-4})$ to the lagged rate; each is bounded below by the floor. Sources: as Table 1; author’s calculations.

Period	Agent		Book rule		Taylor (1993)		Balanced		First diff.
	Mean	RMS	Mean	RMS	Mean	RMS	Mean	RMS	Mean
All quarters	+0.77	1.03	+0.40	0.60	+1.48	2.62	+1.60	2.97	+0.22
Pre-floor, 1997:Q1–2008:Q4	+0.74	0.82	+0.45	0.67	+0.13	1.40	+0.29	1.10	-0.04
First floor spell	+0.03	0.11	+0.23	0.34	+1.04	1.32	+0.28	0.63	+0.19
Second floor spell	+1.11	1.42	+1.09	1.17	+4.07	4.86	+3.81	4.92	+1.00
From 2016:Q1, floor excluded	+1.34	1.49	+0.33	0.46	+3.15	3.73	+4.03	4.78	+0.43
2021:Q1–2023:Q4	+1.96	2.09	+0.52	0.81	+6.12	6.51	+7.24	7.76	+1.75

8.2 The agent frozen at the end of the first floor spell

The comparison above trains the agent on the same record it is compared with. The test of what that hindsight contributes freezes everything at 2015:Q4, the final quarter before liftoff from the first floor spell: the data record truncated to its first 80 rows (28 on the floor), the survey row re-measured on that window (the wedge mean 0.5400 against the full window’s 0.5052), the anchored estimator run from ten catalogued random starts over the box of record, the residual components re-fitted, and the agent selected by the search of Section 7.3 on the truncated blocks with starts through the cutoff and its own catalogued validation starts. The moments of the measurement errors are those known at the cutoff (Section 2). Nothing after 2015:Q4 reaches the frozen agent; it then prescribes a rate at each of the 42 real-time states from 2016:Q1 to 2026:Q2, the state filtered under the frozen blocks.

Table 6 places the frozen estimate beside the full-sample estimate. The record through 2015:Q4 does not identify the Phillips block of the full record: seven of the ten starts converge to one family with λ_1 at 0.693 and a Phillips slope of 0.0172 (standard error 0.0037), and the other three to a flat slope; the full sample’s slope is 3.5 times the frozen one and lies outside two of its standard errors. The frozen policy block holds c_π at the convention’s value, as every estimate does. Trained on that world, the frozen agent’s checkpoint of record stands at step 1,690 of its search, at ρ 0.63, c_π 2.66, c_y 2.92, and c_0 -0.56: responses larger than the agent of record’s (1.23 and 1.34 times its inflation and gap responses), whereas under every earlier construction they had been a fraction of them; the prediction registered with the change to moments known at each decision point, that they would remain a fraction, therefore failed once a later change lowered the agent of record’s responses (Section 2). The checkpoint’s stream ranks eighth of eight on the frozen search’s test set (Section 7.3).

Table 6: The estimate frozen at 2015:Q4 beside the estimate of record. The frozen estimate is the deepest of ten random starts of the anchored estimator on the first 80 rows of the data record, under the box of record with the anchoring weight fixed at one and c_π held at the convention's value of 3.0, which carries no standard error; the full-sample estimate is that of Table 1. Standard errors from the finite-difference Jacobian of the standardized innovations, as in Table 1; a_r is calibrated and carries none. Sources: as Table 1; author's calculations.

Block parameter	Frozen at 2015:Q4		Full sample	
	Estimate	SE	Estimate	SE
IS: lagged gap, a_{lag}	0.8832	0.0736	0.7714	0.0299
IS: expected gap, a_{lead}	0.0868	0.0826	0.1986	0.0343
Phillips: expectation weight, λ_1	0.6927	0.0237	0.9455	0.0414
Phillips: slope, κ	0.0172	0.0037	0.0596	0.0166
Rule: smoothing, ρ	0.7737	0.0305	0.6558	0.0274
Rule: inflation response, c_π	3.0000	—	3.0000	—
Rule: gap response, c_y	1.3111	0.2702	0.8437	0.2787
Rule: intercept, c_0	0.5822	0.2137	3.0008	0.7993
Rule shock persistence, ϕ_u	0.5718	0.0666	0.9586	0.0162
Rule shock sd, σ_η	0.3237	0.0473	0.9474	0.1029
Rows	80		122	

Table 7 gives the result. Over the 42 quarters after the cutoff the frozen agent sits +0.235 points above the FOMC on average (Newey–West standard error 0.138 with 3 lags) with a root-mean-square difference of 0.568, where the agent of record sits +1.297 above; the two agents' rates differ by 1.30 points root-mean-square while moving together (correlation 0.96). The frozen agent holds the floor with the FOMC through 2021, lifts off in 2022:Q1, one quarter before the FOMC's liftoff in 2022:Q2, and tightens ahead of the FOMC through 2023:Q2, by as much as 1.31 points, so that over 2021:Q1–2023:Q4 it lies above the realized rate on 7 of 12 quarters, +0.433 on average against the agent of record's +1.955; from 2023:Q3 it runs with the realized rate or below it, below on 10 of 12 quarters and -0.24 on average, and from 2022:Q2 it is +0.119 from the FOMC. Of the four registered predictions, the sign of the frozen agent's distance held and the three that expected it to resemble the agent of record failed.

Table 7: The frozen agent, the rolling real-time agent, the agent of record, and the book rule against the FOMC after the cutoff. Mean and RMS are the mean and root-mean-square difference between the rule's first-quarter rate and the realized funds rate at the decision point, in percentage points; the frozen agent's rate is computed in the frozen economy at its estimate, the rolling agent's in the economy re-estimated through each decision point with its rule re-trained there, and the other two in the economy of record, each at the real-time state of the quarter. Low and High are the lowest and highest of the mean differences of the rolling windows' eight checkpoints, each stream held fixed across the windows. Sources: as Table 1; author's calculations.

Period	Quarters	Frozen agent		Rolling agent				Full-sample agent		Book rule	
		Mean	RMS	Mean	RMS	Low	High	Mean	RMS	Mean	RMS
2016:Q1–2026:Q2	42	+0.23	0.57	+1.35	1.76	+1.34	+1.35	+1.30	1.48	+0.47	0.66
2016:Q1–2020:Q1	17	+0.39	0.60	+0.48	0.74	+0.48	+0.51	+0.79	0.82	+0.54	0.57
Second floor spell	8	+0.16	0.36	+0.79	1.13	+0.78	+0.80	+1.11	1.42	+1.09	1.17
2021:Q1–2023:Q4	12	+0.43	0.66	+2.15	2.45	+2.09	+2.16	+1.96	2.09	+0.52	0.81
2022:Q2–2026:Q2	17	+0.12	0.61	+2.49	2.55	+2.45	+2.48	+1.89	1.94	+0.11	0.32

Figure 9 draws the three paths. At the pandemic's first quarter both agents cut short of the floor the FOMC reached and reach it the quarter after; the frozen agent then holds the floor with the FOMC, leaves it a quarter earlier, and from 2023:Q3 runs with the realized rate, where the agent of record runs +1.89 above it from 2022:Q2.

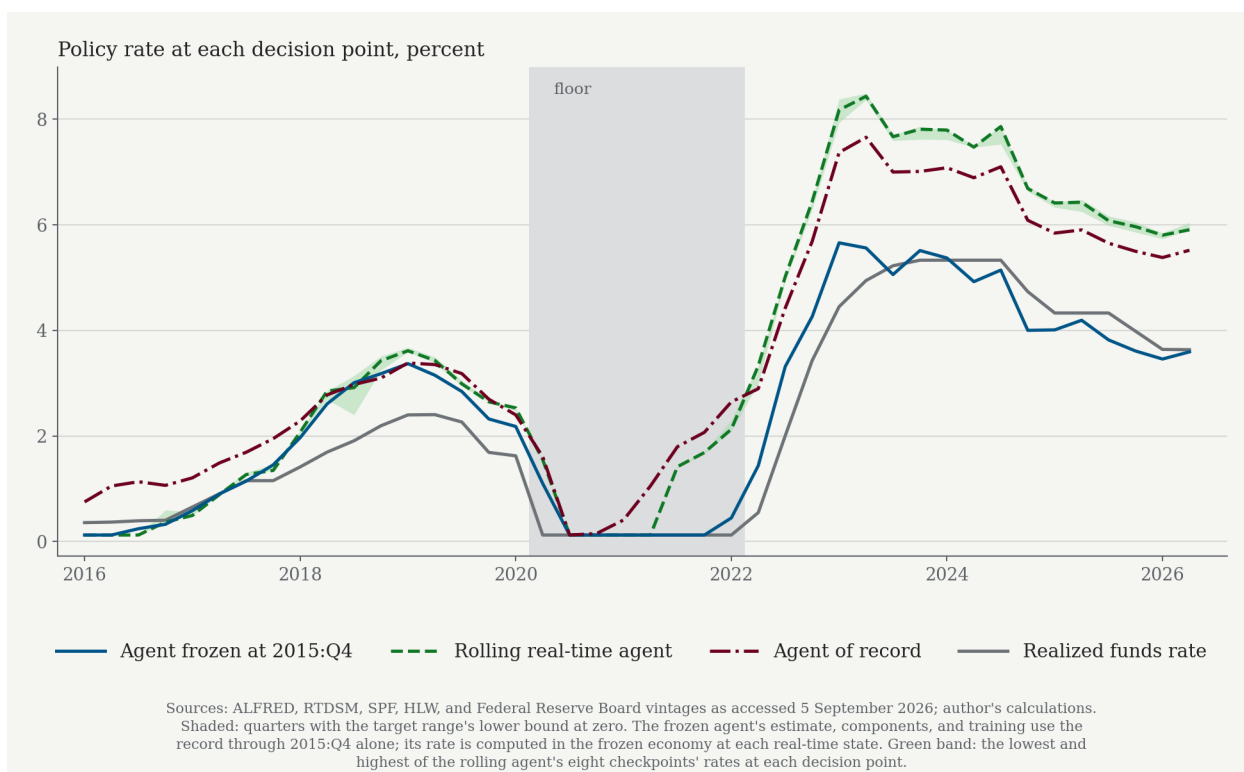


Figure 9: The agent frozen at 2015:Q4, the rolling real-time agent, and the agent of record against the FOMC at the 42 real-time decision points from 2016:Q1 to 2026:Q2. The deep blue line is the frozen agent's first-quarter policy rate in the frozen economy at its estimate; the maroon dash-dot line is the agent of record's rate in the economy of record; the gray line is the realized funds rate at the decision point, at the model's floor where the target range's lower bound was zero (the shaded spell). The frozen agent's mean difference from the realized rate is +0.235 percentage points (Newey–West standard error 0.138) against the agent of record's +1.297 on the same quarters. The green dashed line is the rolling agent's rate, re-estimated and re-trained at each decision point on the record through that point, and the green band the range of its eight checkpoints' rates at that point. Sources: ALFRED, RTDSM, SPF, HLW, and Federal Reserve Board vintages as accessed 5 September 2026; author's calculations.

The reading is that the distance of the agent of record from the FOMC is, in large part, learned from the record after the cutoff. The Phillips slope that puts the optimal rate above the FOMC's is identified by the inflation of 2021–2023, not by the twenty years that end at the cutoff; an agent that knew only what could be known at the end of the first floor spell prescribes rates within a quarter point of the FOMC's on average thereafter. The full-sample comparison stands as what the loss of record and the estimated world would have set with the whole record in hand; the frozen comparison states what the same construction would have set with the record of the day.

The frozen agent is a filter stopped at one date and run forward without updating; the agent of record is the smoothed estimate, every quarter's prescription conditioned on the whole record. Between them stands the filtered one: at each of the 82 quarterly decision points from 2006:Q1 to 2026:Q2 the blocks are re-estimated on the record through that point, under the estimate-of-record rule with 10 random starts over the box and the previous point's optimum as one more, the rule is re-trained on those blocks under the stopping rule of Section 7.3, and its first-quarter rate

is taken at that point's real-time state filtered under those blocks. The neutral-rate path is real-time from 2016:Q1, each quarter at the vintage the panel carries at its decision point and the last vintage held through the pandemic suspension, and the ex post path before that, since no real-time vintage exists; the line is real-time in structure, state, and filter throughout and in the neutral rate from 2016:Q1. The modeling choices are today's, as they are for every agent in the paper; only the conditioning set is real-time. The learned rule cannot start earlier than 2006:Q1, forty quarters into the record, and the first window is chosen so that the financial crisis and the first floor spell fall inside the rolling line.

Table 7 and Figure 9 carry the rolling agent beside the other two over the frozen window. Over the 42 quarters from 2016:Q1 its mean difference from the FOMC is +1.35 points (Newey–West standard error 0.322, 3 lags), against the frozen agent's +0.235 and the agent of record's +1.297; it is +0.48 over 2016–2020 and +2.15 in the windows whose record includes 2021, so the distance opens as the record takes in the inflation. Over the 40 quarters from 2006:Q1 to 2015:Q4, through the crisis and the first floor spell, the rolling agent's mean difference is +0.02 (standard error 0.044; RMS 0.18): the record of the day prescribed what the FOMC did. Three predictions were registered for the rolling line, each against the value an earlier construction of it had given. Two held: the mean difference through 2015 lies within one standard error of the earlier value, and from 2016 the difference keeps its sign and is larger once the record includes 2021. One failed: the window estimate's λ_1 was to reach its structural limit within eight quarters of 2022:Q1, where the earlier construction first found it there, and it reaches the limit in no window: it rises from 0.69 at 2021:Q2 to 0.99 at 2021:Q4 as the inflation enters the record and settles below the limit, as the estimate of record's does; at the final window the optimum stands at 0.1335 with λ_1 at 0.93 and κ at 0.0498 under the real-time neutral-rate path. An expectation carried from the earlier construction fails as well: the rolling agent's mean difference from 2016 was to lie between the frozen agent's and the agent of record's, and it lies 0.06 above the agent of record's, a distance inside the rolling line's own standard error. In three windows one random start could not be followed by the estimator, the filter's convergence set changing under a finite-difference perturbation; it is recorded as failed and the window's optimum comes from the starts that ran. No training reached the step cap.

Each rolling window is also a search of the kind Section 7.3 describes, eight streams and a selected checkpoint, so the 82 windows read the selection where the paper's other searches cannot, on many estimates. Three readings were registered. The first is whether a window's validation set ranks its eight checkpoints as its test set does; three predictions were registered for it, the three among those registered when the validation set was replaced (Section 7.3) that concern the rolling windows: a positive median Spearman correlation of the checkpoints' validation and test losses across the windows, a positive correlation in 42 or more of the 82 windows, and the selected stream in the better half on the test set in 42 or more. The validation set does not: across the windows the median Spearman correlation of the checkpoints' validation and test losses is -0.21 , the correlation is positive in 30 windows, and the selected stream ranks in the better half on the test set in 36, so all three failed. The second is whether a larger selection set would do better, read against a fresh held-out set of the same size drawn for every window after its search, with no retraining; four predictions were registered for it: a positive median correlation, across the windows, of the pooled criterion (the mean of a checkpoint's validation and test losses) with the checkpoint's loss on the fresh set; the same for the test loss alone, the control; a selection on the pooled criterion in the better half on the fresh set in 42 or more windows; and, among the windows where the pooled selection and the construction's own, made on the validation set, differ, the lower fresh loss for the pooled selection in more than half. The mean of a checkpoint's validation and test losses correlates with its loss on the fresh set at a median of +0.30, and the test loss alone at +0.40, where the validation loss

alone stands at -0.08 : the checkpoints differ by more than the sampling of a set of this size, and the validation loss, the lowest of a stream's readings, does not carry to a fresh set. A selection on the pooled criterion ranks in the better half on the fresh set in 43 windows, and in the 62 windows where it differs from the construction's selection it has the lower fresh loss in 34. All four held, the two on the selection by margins that an even chance gives with probability 0.37 and 0.26; the construction's own selection ranks in the better half on the fresh set in 41 of the 82 windows, as a draw among the eight would. What the selection decides is small: at the median window the eight checkpoints' fresh losses lie within 0.56 percent of one another, and the selected checkpoint's exceeds the lowest by 0.16 percent. The third reading is whether the rolling line depends on the selection. At each window every one of the eight checkpoints prescribes a rate at the window's state, and the statistics above are taken on each stream held fixed across the windows. On all eight the mean difference through 2015 is within half a point ($+0.01$ to $+0.02$), the difference over 2016–2020 ($+0.48$ to $+0.51$) is smaller than the difference from 2021 ($+2.12$ to $+2.14$), and the mean from 2016 ($+1.34$ to $+1.35$) lies above the agent of record's. The lowest and highest of the eight rates are a median of 0.00 points apart through 2015, 0.06 over 2016–2020, and 0.13 from 2021, and 0.74 at the widest window, 2018:Q3; Figure 9 shades the band and Table 7 gives the lowest and highest of the streams' means. The rolling line's reading does not rest on which stream the selection names, and neither does the expectation that failed.

What the rolling line adds to the reading is the timing. A learner who kept learning stayed near the FOMC until the record contained the inflation, and moved away from it afterward, as far as the agent of record stands with the whole record in hand; the FOMC's path stayed near the frozen line throughout. The frozen result shows the FOMC's decisions consistent with a rule that did not learn from 2016 on; the rolling result shows that a rule which did learn would have departed from them once the evidence for the steeper block was in hand, and that before then it stood where the frozen rule stands, $+0.48$ above the FOMC over 2016–2020 against the frozen agent's $+0.38$: there was little in the record of the day to depart on. Neither says what the FOMC learned; both describe its decisions against rules that condition on stated records.

8.3 Counterfactual outcomes

The comparisons above are of rates. What the rules would have produced is a second question, and Table 8 answers it under a construction whose limits are stated first. The structural residuals of the IS and Phillips blocks are recovered from the realized record at the estimate: the filter's residual states each quarter, the innovations the model needs to reproduce the observed inflation and rate given the filtered gap. Fed back under the estimated policy block with its persistent disturbance, they reproduce the observed inflation to a root-mean-square error of 0.009 (largest 0.046) and the rate to 0.0102 on the 86 quarters where the rate row is observed off the floor. The economy, the blocks alone as in the comparison of Section 8.1, is then simulated from 1997:Q1 under the agent's rule with those residuals as each quarter's shocks, expectations model-consistent under the rule and future shocks zero, the state advancing on the model's own outputs; the persistent disturbance of the estimated policy block, which is the FOMC's deviation from its own reaction function, is not fed. Two conditions bound the reading: the private sector's expectations are model-consistent under the simulated rule throughout, and the residuals are treated as invariant to the rule, which is the condition Lucas named;¹³ and the gap on the realized path is the filter's own fitted gap, because

¹³Robert E. Lucas Jr., "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.

the true gap is latent, so the two models read the realized record differently. Neither path is an outcome the FOMC could have chosen.

Under the agent of record in its economy, four-quarter core inflation over 2021:Q1–2023:Q4 averages 2.25 against the realized 4.29, the gap -0.07 against 2.72, and the rate 2.66 against 2.21: the rule contains the inflation with a policy rate only somewhat above the realized one, the containment running through expectations consistent with a rule whose inflation response is 2.17. The book loss over the record falls from 989.3 to 261.6. Under the frozen agent in the frozen economy, whose reading of the same record puts the 2021–2023 gap at -0.19 rather than 2.72, inflation over the episode averages 4.42 against 4.29 and the loss since the cutoff is 148.0 against the realized 153.0 under that reading: the record-of-the-day rule leaves the episode's inflation 0.14 above the realized where the rule of record lowers it by 2.03, and lowers the loss since the cutoff only slightly. Before the cutoff, where the frozen rule was trained, its loss is 117.7 against the realized 620.0. The six registered predictions held; the frozen counterfactual's loss since the cutoff lies between the rule of record's and the realized path's under both models' readings of the gap, 343.5 and 153.0.

The containment's dependence on the expectations block is measured by re-estimating the blocks with λ_1 held at the book's 0.70 and everything else as in the estimate of record, from the estimate of record and from run B as starts, and running the rule of record, unchanged, in that economy with the residuals recovered at that estimate. From the estimate of record the search converges in 12 iterations to an objective of 0.1820 against the record's 0.1445, about 31 units of log-likelihood worse, so the record's λ_1 is the optimum and the book's value is rejected, with a Phillips slope of 0.0310 (standard error 0.0095) against the record's 0.0596 and a policy smoothing of 0.707; from run B as the start the search reaches a nearby optimum a little deeper, at 0.1807, and the sensitivity stays with the first start's, as registered. In the $\lambda_1 = 0.70$ economy the rule of record holds 2021–2023 inflation at 2.29 against the realized 4.29, above the 2.25 of the record's economy and below the realized by 1.99 points, 0.98 of the record's reduction, with the loss over the record at 217.3 against the realized 914.4 under that model's reading of the gap. The four registered predictions held: the containment runs through the expected path in both economies and does not rest on a Phillips block that leans as far forward as the record's.

The realized rows differ only in the gap, and a third reading of it comes from outside the model: potential output as the Congressional Budget Office (CBO) projected it in its 27 February 2026 vintage, against real GDP at the 15 August 2026 vintage, both from the store of Section 2. Under that reading the 2021–2023 gap averages 0.93, between the frozen filter's -0.19 and the record filter's 2.72, and its 1997–2015 mean is -0.76 ; the realized book loss is 594.4, between the agent of record's 261.6 and the record filter's 989.3. The three registered predictions held. The CBO reading is a third construction of the gap and not a benchmark for the two filters; what it adds is that the ordering of the agent of record's counterfactual loss against the realized loss holds under a gap the model did not filter.

Table 8: Counterfactual outcomes under the agent of record and under the frozen agent, each in its own economy with the residuals recovered at its own estimate, and under the agent of record in the economy re-estimated with λ_1 held at 0.70, beside the realized path as each model reads it. Inflation is four-quarter core PCE inflation in percent, the gap in percent of potential, the rate the funds rate in percent, each averaged over 2021:Q1–2023:Q4; the book loss is the loss of record’s quarterly loss summed over 1997:Q1–2015:Q4, 2016:Q1–2026:Q2, and the record, in the loss’s units. The realized rows share the observed inflation and rate and differ in the gap, which is each model’s filtered gap. The final row reads the gap as the log ratio of real GDP to CBO potential output, in percent. Sources: as Table 1; CBO potential output (ALFRED, series GDPPOT, vintage of 27 February 2026) and real GDP (RTDSM, ROUTPUT, vintage of 15 August 2026); author’s calculations.

Path	2021:Q1–2023:Q4, means			Book loss		
	Inflation	Gap	Rate	1997–2015	2016–2026	Total
Realized, gap read by the model of record	4.29	+2.72	2.21	645.8	343.5	989.3
Agent of record, its economy	2.25	-0.07	2.66	219.3	42.3	261.6
Realized, gap read by the frozen model	4.29	-0.19	2.21	620.0	153.0	773.0
Frozen agent, the frozen economy	4.42	-1.34	3.47	117.7	148.0	265.7
Realized, gap read by the $\lambda_1 = 0.70$ model	4.29	+1.93	2.21	670.9	243.5	914.4
Agent of record, the $\lambda_1 = 0.70$ economy	2.29	-0.47	2.06	172.4	44.9	217.3
Realized, gap from CBO potential output	4.29	+0.93	2.21	370.9	223.5	594.4

9 Sensitivity and Limitations

The loss weights are the book’s and were not estimated, and Table 9 measures what they decide. The search of the agent of record was repeated under four variants of the weights, β halved and raised by half and γ halved and raised by half (the first of these is the book’s own robustness check, its DM2, which halves the same weight and which the book reports as lying almost on top of DM1), with α held at 1.0 as the normalization and with the draws, the starts, the validation and test sets, and the random streams of the record unchanged, so that a variant differs from the agent of record in the weights alone; each variant’s search selects its own checkpoint under its own weights, and each is evaluated as the record is. The coefficients do not all move in the direction the weight names. The smoothing does: ρ rises from 0.595 to 0.673 across γ , the record’s 0.644 between them. The gap response does not: c_y is 2.007 at $\beta = 0.5$ and 1.562 at $\beta = 1.5$, with the record’s 2.176 above both, against the registered prediction. Across the grid c_y runs from 1.56 to 2.42, and across the five searches at the weights of record (Table 10) from 1.99 to 2.60: the grid moves the gap response by about as much as the choice among searches does, and at this grid’s spacing its effect on that coefficient is not resolved. The prescriptions move less than the coefficients. The full-record mean difference to the realized rate runs from +0.614 at $\beta = 0.5$ to +0.902 at $\gamma = 0.25$, every variant within 0.16 points of the record’s +0.768 and every variant above the FOMC by at least 5.5 Newey–West standard errors. Every variant beats the book rule in the exact economy at the estimate of record under its own weights. Under the weights of record the variant with the rate-change weight halved has a loss there of 5.907, 0.0154 below the agent of record’s 5.923 and within the 0.0949 by which the five searches’ losses spread at the weights of record: the agent of record is the optimum of its own objective to within the search’s own dispersion, and not closer. Closeness to the realized path is not a criterion on the weights, the objective being the loss as stated, and the record does discriminate among them on it: the variant nearest the FOMC is the one with the rate-change

weight raised by half, with an RMS difference of 0.797 against the record's 1.026, and the farthest the one with the rate-change weight halved, at 1.238. Of the five registered predictions, the three on the prescriptions, on the book rule, and on the record's own objective held, and the two on the gap response's direction and on the record's silence among the weights failed. A loss that weights the gap or the rate change differently moves the optimal rule nearer the FOMC or farther from it by tenths of a point; within the grid it does not close the distance.

Table 9: The agent of record and four agents trained under other loss weights, everything else unchanged. The weights are (α, β, γ) with $\alpha = 1.0$; the record's are $\beta = 1.0$ and $\gamma = 0.50$, and each variant changes the one it names. Stream and step are the selected stream of the row's search over eight and the step of its checkpoint, the horizon being 2,000 steps; ρ , c_π , and c_y are the smoothing, inflation, and gap responses; the mean difference is the agent's rate less the realized funds rate over the 118 decision points, in percentage points, with its Newey–West (NW) standard error and the RMS difference. Loss, own is the book's loss in the exact economy at the estimate of record under the row's own weights, where the book rule's under the same weights runs from 5.424 to 14.464 across the variants; loss, record is the same loss under the weights of record, where the book rule's is 9.944. Sources: as Table 1; author's calculations.

Weights	Stream	Step	ρ	c_π	c_y	Mean diff.	NW SE	RMS	Loss, own	Loss, record
Record	8	450	0.644	2.165	2.176	+0.768	0.128	1.026	5.923	5.923
$\beta = 0.5$	8	440	0.712	1.955	2.007	+0.614	0.095	0.806	3.402	5.976
$\beta = 1.5$	7	240	0.624	1.905	1.562	+0.709	0.102	0.899	8.749	6.151
$\gamma = 0.25$	8	520	0.595	2.329	2.419	+0.902	0.162	1.238	5.663	5.907
$\gamma = 0.75$	7	260	0.673	1.884	1.569	+0.627	0.089	0.797	6.298	6.102

The training itself is one more source of variation, and Table 10 measures it. The search of the agent of record was repeated under four further sets of catalogued streams, everything else unchanged, each a search over eight streams at the horizon, none extended, and each evaluated as the record is. The coefficients move: c_π spans 2.07 to 2.56 and c_y 1.99 to 2.60 across the five, with ρ between 0.64 and 0.67. The prescriptions do not: the full-sample mean difference to the realized rate runs from +0.733 to +0.789, a range of 0.056 points, with Newey–West standard errors of 0.116 to 0.142, and every member's loss in the exact economy at the estimate of record lies between 5.88 and 5.97 against the book rule's 9.94; the three registered predictions held. The distance from the FOMC is a property of the estimated economy and the loss, not of one search; the rule's coefficients are resolved to the spread across searches and no further, and the paper reads them to that precision.

Table 10: The agent of record and four agents trained under other random streams, everything else unchanged. Stream and step are the selected stream of each search over eight and the step of its checkpoint, the horizon being 2,000 steps; test rank is the selected stream’s rank of eight on the search’s held-out test set; ρ , c_π , and c_y are the smoothing, inflation, and gap responses; the mean difference is the agent’s rate less the realized funds rate over the 118 decision points, in percentage points, with its Newey–West (NW) standard error; the loss is the book’s, in the exact economy at the estimate of record, where the book rule’s is 9.94. Sources: as Table 1; author’s calculations.

Agent	Stream	Step	Test rank	ρ	c_π	c_y	Mean diff.	NW SE	Loss
Agent of record	8	450	3	0.644	2.165	2.176	+0.768	0.128	5.923
Member 1	2	310	3	0.647	2.073	1.986	+0.733	0.116	5.971
Member 2	8	560	3	0.655	2.355	2.421	+0.789	0.138	5.888
Member 3	8	620	1	0.670	2.565	2.599	+0.787	0.142	5.876
Member 4	5	680	3	0.650	2.121	2.235	+0.765	0.129	5.917

Eight limitations bound the reading. The IS residual carries structure beyond the floor-aware null under the estimate of record, so the IS block is incomplete in a way the components measure but the training environment does not use. The estimate of record’s Phillips block leans nearly all the way forward, λ_1 at 0.946, so inflation in the estimated economy is carried largely by model-consistent expectations of it, which is the channel through which the counterfactuals of Section 8.3 contain it; the containment survives a block with the book’s weight on the lagged term at 0.98 of the reduction, so the reading rests on the channel and not on the weight’s estimated value. The loss carries the target-deviation form of the 2018 book, while the FOMC’s own framework changed in 2020 to a form with history dependence, so the agent’s behavior through 2021–2023 is a statement about that loss as much as about the agent. And the oracle’s floor spell is shorter than the record’s first spell, so the certifications bound the estimator’s behavior on floor spells of the oracle’s length, not of 28 quarters. The fifth bounds the comparison itself: the estimated policy block is the FOMC’s reaction function as the record carries it, and the agent’s rate at each state is computed under expectations consistent with the agent’s rule, so the comparison holds the private sector’s expectations formation fixed across the two policies, the condition of Section 8.3. The sixth is the convention. The scale of the estimated policy block is fixed at $c_\pi = 3.0$ by registration and not estimated; along the profile the objective falls to 0.1349 at 8, about 8 units of log-likelihood deeper than at the convention, and an agent trained on the estimate at the box’s bound sat farther above the FOMC over 2016 on than one trained at 3 (0.97 against 0.57 points, under the construction before the convention). What the agent learns does not turn on the scale; how much above the FOMC it sits does, and the convention is a choice the paper states and does not estimate. The seventh is the neutral rate. The forward solve at a decision point reads the path of Section 2 from that quarter on, so the states of Section 8 are real-time in their observations and not in the neutral rate the expectations are formed on, except for the rolling agent of Section 8.2, whose neutral rate is real-time from 2016:Q1 and ex post before; the path runs from 0.57 to 3.90 percent over the window, and a real-time construction of it is a change of record the paper does not make. The eighth is the selection among training streams. The checkpoint of record is the best of eight on a validation set of eight quarters, and on the held-out test set the eight streams’ ranking does not follow the validation ranking in most of the paper’s searches (Section 7.3); the prescriptions the paper reports are those of the selected checkpoint, and the spread across searches of Table 10 is the measure of what a different selection would have prescribed.

10 Conclusion

The frontier the book describes is a policy that minimizes a loss against a model. This paper asked what that policy is when the model must be learned, and answered by measuring the four learners the question contains. The economy's agents, in the macro rows, do not learn in any of the three senses the literature proposes: a fixed anchor, a compounded discount, and an anchor learned from experience each fold back to model-consistent expectations on the full window, under an estimator certified to find each of them when it is there; the long-run survey does learn, at a measured speed, and is more anchored than a learner's mean of experience. The econometrician learns a likelihood that does not pin the scale of the policy block, fixed here by a stated convention, and whose deepest basin the record's own catalogued points find where a random search does not; the reading of the world changes between the basins, from a Phillips block that leans on lagged inflation with a flat slope to a nearly forward-looking block with a resolved slope and a persistent policy shock. The correction learns structure that a floor spell manufactures at the true parameters, and against a null that knows this it carries no signal in the inflation block. The policymaker learns a rule that reproduces the risk-avoidance timing of the book's optimal path at the floor, beats the simple rule in its own world on every draw from the estimate's uncertainty and in the book's calibration, and sat above the FOMC across the record except when both were at the floor; frozen before liftoff, the same construction sits within a quarter point of the FOMC thereafter, because the record of the day does not identify the Phillips slope that the later inflation identifies.

The bound on what the fourth learner can exploit is the estimate's own uncertainty: a policy trained across the standard-error band can exploit what the estimate supports and no more, and the agent that was trained on the correction instead beat the simple rule on fewer than half of those draws. Optimal monetary policy when the economy must be learned is therefore not a single path but a rule whose optimality is conditional on four certified learners, and what it says about the FOMC, that the loss of record and the estimated world would have set rates higher throughout the record except at the first floor, is stated as a property of that conditional object rather than a judgment of the FOMC, one that the weights of the loss move within the grid by tenths of a point and do not close, and the frozen comparison bounds it: what the construction says about the FOMC depends on what the record had taught it by the decision date, and so do the outcomes it would have produced, which the rule of record improves on the whole record and the rule of the day, on its own reading of the record, improves chiefly where it was trained. What would change the reading is stated at each step: an expectations construction that varies across regimes, a loss with the history dependence of the FOMC's 2020 framework, a scale of the policy block the record pins rather than a convention fixes, and a selection among training streams that its own test set confirms.

A The Oracle Conditions

Each condition simulates the economy at known parameters, applies the estimator, and records whether registered predictions on parameter recovery held. Conditions O1 through O14 certify the estimator on pre-floor trajectories, its treatment of measurement error, revision, and missing rows, and the anchored expectations block; O15 through O18 carry a scripted floor spell and, from O16, the record's missing-row pattern, and certify the anchored, cognitively discounting, and learned-anchor blocks in turn. Under the floor spell the anchored weight is recovered with standard errors of 0.039 to 0.048 when the staff rows are absent on the final 22 quarters (O16; with every row present, O15, 0.037 to 0.046), the discount with standard errors of 0.022 to 0.033, and the learned gain with standard errors below 0.0028. Each condition's predictions were stated before its estimation units ran, and its outcomes are reported against them.

Bibliography

- Adrian, Tobias, Richard K. Crump, and Emanuel Moench. "Pricing the Term Structure with Linear Regressions." *Journal of Financial Economics* 110, no. 1 (2013): 110–138.
- Adrian, Tobias, Douglas Laxton, and Maurice Obstfeld, eds. *Advancing the Frontiers of Monetary Policy*. Washington, DC: International Monetary Fund, 2018.
- Alichi, Ali, Kevin Clinton, Charles Freedman, Michel Juillard, Ondra Kamenik, Douglas Laxton, Jarkko Turunen, and Hou Wang. "Avoiding Dark Corners: A Robust Monetary Policy Framework for the United States." IMF Working Paper 15/134. Washington, DC: International Monetary Fund, 2015.
- Blanchard, Olivier. "Where Danger Lurks." *Finance & Development* 51, no. 3 (2014): 28–31.
- Brainard, William C. "Uncertainty and the Effectiveness of Policy." *American Economic Review* 57, no. 2 (1967): 411–425.
- Brayton, Flint, Thomas Laubach, and David Reifschneider. "Optimal-Control Monetary Policy in the FRB/US Model." FEDS Notes. Washington, DC: Board of Governors of the Federal Reserve System, November 21, 2014.
- Chen, Mingli, Rama Cont, Andreas Joseph, Michael Kumhof, Xinlei Pan, Wei Xiong, and Xuan Zhou. "Deep Reinforcement Learning in a Monetary Model." Staff Working Paper 1,142. London: Bank of England, 2025.
- Edge, Rochelle M., Thomas Laubach, and John C. Williams. "Welfare-Maximizing Monetary Policy under Parameter Uncertainty." *Journal of Applied Econometrics* 25, no. 1 (2010): 129–143.
- Evans, George W., and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press, 2001.
- Gabaix, Xavier. "A Behavioral New Keynesian Model." *American Economic Review* 110, no. 8 (2020): 2271–2327.
- Ha, David, and Jürgen Schmidhuber. "Recurrent World Models Facilitate Policy Evolution." In *Advances in Neural Information Processing Systems* 31 (2018): 2450–2462.
- Hansen, Lars Peter, and Thomas J. Sargent. *Robustness*. Princeton, NJ: Princeton University Press, 2008.
- Hinterlang, Natascha, and Alina Tänzler. "Optimal Monetary Policy Using Reinforcement Learning." Discussion Paper 51/2021. Frankfurt am Main: Deutsche Bundesbank, 2021.
- Holston, Kathryn, Thomas Laubach, and John C. Williams. "Measuring the Natural Rate of Interest: International Trends and Determinants." *Journal of International Economics* 108 (2017): S59–S75.
- Krishnamurthy, Arvind, and Miguel Chumbo. "Disappearing Dollar Convenience?" NBER Working Paper 35742. Cambridge, MA: National Bureau of Economic Research, 2026.
- Lucas, Robert E., Jr. "Econometric Policy Evaluation: A Critique." *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.
- Orphanides, Athanasios. "Monetary Policy Rules Based on Real-Time Data." *American Economic Review* 91, no. 4 (2001): 964–985.
- Primiceri, Giorgio E. "Why Inflation Rose and Fell: Policy-Makers' Beliefs and U.S. Postwar Stabilization Policy." *Quarterly Journal of Economics* 121, no. 3 (2006): 867–901.

- Sargent, Thomas J. *The Conquest of American Inflation*. Princeton, NJ: Princeton University Press, 1999.
- Sutton, Richard S. “Dyna, an Integrated Architecture for Learning, Planning, and Reacting.” *ACM SIGART Bulletin* 2, no. 4 (1991): 160–163.
- Taylor, John B. “Discretion versus Policy Rules in Practice.” *Carnegie-Rochester Conference Series on Public Policy* 39 (1993): 195–214.
- Taylor, John B., and John C. Williams. “Simple and Robust Rules for Monetary Policy.” In *Handbook of Monetary Economics*, vol. 3B, edited by Benjamin M. Friedman and Michael Woodford, 829–859. Amsterdam: Elsevier, 2010.
- Wieland, Volker. “Monetary Policy, Parameter Uncertainty and Optimal Learning.” *Journal of Monetary Economics* 46, no. 1 (2000): 199–228.
- Wu, Jing Cynthia, and Fan Dora Xia. “Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound.” *Journal of Money, Credit and Banking* 48, no. 2–3 (2016): 253–291.