Unspanned Reserve Adequacy and the Term Premium in Federal Funds Futures

Zong-Wei Yeh* Dongjie Fang†

October 18, 2026

Abstract

The floor system separates the size of the Federal Reserve's balance sheet from control of the policy rate. We ask whether the separation extends to the price of risk: whether the quantity of reserves is priced in the market for the policy rate. In federal funds futures, reserve adequacy, measured as reserve balances relative to GDP, is almost invisible in the current curve yet forecasts the curve's evolution, in sample and out of sample. An arbitrage-free affine model in which reserve adequacy drives the physical but not the risk-neutral dynamics attributes about four-fifths of the variance of the twelve-month term premium to it: a one-standard-deviation decline in reserve adequacy raises the premium by 41 basis points, and the premium widens as reserves run off and narrows as they are added. The overnight market shows why. A given reserve drain moves the funds rate only when reserves are scarce, which is what gives reserve adequacy its importance; the futures curve ignores the drains and responds to the persistent state that governs their effect. The compensation embedded in federal funds futures varies with this state, rising as reserve balances approach the steep region of reserve demand.

Keywords: Monetary Policy, Reserve Adequacy, Federal Funds Futures, Term Premium, Unspanned Macro Risk.

JEL codes: G12, E43, E52, G21.

*Department of Finance, National Chung Cheng University. Email: zweyeh@ccu.edu.tw.

†Kate Tiedemann School of Business and Finance, University of South Florida. Email: dfang@usf.edu.

1 Introduction

The floor system rests on a separation. The size of the Federal Reserve's balance sheet is set by asset purchases and their runoff, the policy rate is set by administered rates, and as long as reserves are ample the two can be moved independently: the effective funds rate stays where interest on reserves puts it whatever the quantity of reserves. For the level of the overnight rate the separation holds by design, and it held until reserves became scarce in September 2019, when the overnight rate jumped.[1] A claim on the future overnight rate, however, carries two things, the expected path of the policy rate and the compensation investors require for the risk around that path, and the separation speaks only to the first. If a smaller balance sheet moves the system toward the steep region of reserve demand, where a given reserve drain moves the funds rate (Afonso et al., 2025), then the risk around the path grows as reserves fall even while the rate itself stands still. This paper asks whether the separation extends to the price of that risk: does the quantity of reserves carry a premium in the market for the policy rate itself?[2]

Answering it requires separating that compensation from the expected policy path, which is what a term structure model does, and a market whose prices are claims on the policy rate, which federal funds futures are. The feature of the data that makes the separation possible is the one the floor system creates. Reserve adequacy, measured as reserve balances over GDP, is almost absent from the current futures curve, which explains about one percent of its variation, and yet it forecasts how the curve moves. A state variable with these two properties is unspanned in the sense of Joslin et al. (2014), and an affine model can exclude it from the pricing map while letting it shape the compensation embedded in longer-dated contracts. We use that model to measure the price of reserve

[1] Overnight repo rates rose by several percentage points on 17 September 2019, the effective funds rate breached the target range, and the Fed resumed reserve-adding operations within days (Afonso et al., 2020; Anbil et al., 2021; Copeland et al., 2025). Reserves at the time were 6.5 percent of GDP, a level that had looked comfortable in the aggregate.

[2] The separation is itself under discussion. Proposals to shrink the balance sheet substantially while keeping the rate tool unchanged (Warsh, 2025), and the Reserve Bank presidents' defense of the ample-reserves framework (Logan, 2026), both turn on how far reserves can fall before the two instruments interact.

scarcity, and we use the overnight market to explain why the price exists.

We measure reserve adequacy as reserve balances scaled by nominal GDP, the yardstick Federal Reserve officials use when they discuss where the ample region ends: Governor Waller places that boundary near nine percent of GDP (Waller, 2025). The measure is built in real time from vintage data, so that nothing a forecaster could not have known at the time enters it. Over the 2010–2025 sample it runs from 6.5 percent of GDP in the week of the September 2019 repo spike to 18.8 percent at the height of pandemic asset purchases, and it traverses two balance-sheet runoffs, in 2017–2019 and from 2022. The market is CME 30-Day Federal Funds futures, the oldest and deepest market for the overnight policy rate and the one in which money-market risk premia have long been studied (Piazzesi and Swanson, 2008; Hamilton and Okimoto, 2011; Cieślak, 2018); we work with the first twelve monthly contracts, the maturities over which balance-sheet policy operates and over which the curve is free of the duration risk that dominates Treasury yields.[3]

Three findings make the case. The first is that reserve adequacy is unspanned and yet forecasts the curve. The leading principal components of the futures ladder leave essentially all of its innovation variation unexplained, and the result does not depend on how reserves are scaled. In one-week-ahead regressions of the kind Heath (2019) uses to test spanning in oil futures, reserve adequacy predicts changes in both the level and the slope of the curve over and above the curve itself and over and above the payroll, credit-spread, and volatility predictors this literature uses. Over longer horizons the associated excess returns trace the profile the economics implies: smallest at one quarter and building through five, with a loading that is negative throughout, more reserves today, lower excess returns ahead. The profile holds out of sample. In the forecasting exercise of Joslin et al. (2014), with parameters estimated only on data available at each forecast date,

[3]Repo-based SOFR futures begin only in 2018 and clear our liquidity screen only from March 2020, so they cannot cover the first runoff or September 2019. On the window where the two markets overlap, an independently constructed one-month SOFR panel gives a reserve-adequacy slope within a fraction of a standard error of the funds-futures estimate (Appendix Table E.6), which argues against an artifact specific to the funds-futures contract, though the overlap is too short to support an independent estimate of magnitude.

adding reserve adequacy to the curve improves pseudo-out-of-sample forecasts of two- to four-quarter excess returns by a quarter to a third, by the test of Clark and West (2007) at conventional significance levels, and the gain survives when the forecasts made in 2023 and 2024 are set aside. The standard return predictors, added on their own, improve nothing.

The second finding is the size of the compensation. We build an arbitrage-free affine model of the money-market curve in which reserve adequacy enters the physical dynamics and is excluded from the price, following Joslin et al. (2014), and adapt it to the settlement conventions of monthly-average contracts. Three curve factors price the twelve maturities to about two basis points. Reserve adequacy accounts for about four-fifths of the model-implied variance of the twelve-month term premium, and the restriction that would set this component to zero, that reserve conditions are absent from the physical forecast of the curve, is rejected. A one-standard-deviation decline in reserve adequacy raises the twelve-month premium by 41 basis points, and the component builds with the contract horizon, from a few basis points at one quarter to the full amount at four, because a persistent state that enters the forecast but not the price contributes a premium that accumulates as the forecast lengthens. Evaluated along the realized path of reserves, the component is widest in the weeks around the September 2019 repo spike, narrowest in the reserve abundance of 2021, and widens again in the second runoff. Pricing reserve adequacy as a spanned factor instead leaves the premium essentially unchanged and gives the factor a curve loading of close to zero, so the compensation is a property of the physical forecast rather than of the current cross-section.

The third finding is the mechanism, and it comes from the overnight market itself. Following Hamilton (1997), we identify reserve drains from movements in the Treasury's balance at the Fed and in overnight reverse repurchase balances, which withdraw reserves one for one and are driven by the tax calendar and money-fund allocation rather than by the Fed's response to funding conditions. When reserves exceed ten percent of GDP, a drain of one percent of GDP moves the spread of the funds rate over the administered

rate by a third of a basis point or less. When they are below, the same drain raises the spread by about two basis points, and the slope is confined to the lowest quintile of reserve adequacy. This is the steepening of reserve demand that Afonso et al. (2025) estimate directly, here recovered with the paper's own state variable. The same drains, on the tax dates and month-ends when they are largest, move the spot rate more when reserves are scarce, while no maturity of the futures ladder responds more in the scarce state; and when the recent change in reserves is added to the level in the one-week forecast, the level does the forecasting and the flow adds little. The curve does not respond to the drains. What its subsequent movements and the model-implied premium reflect is the slow-moving state that decides how much the next drain will matter, and the premium is the compensation for sitting in the state in which it matters more.

The compensation we measure is the price the market attaches to the state of reserve adequacy, denominated in the units of the policy discussion: basis points of twelve-month term premium per log point of reserves relative to GDP. It is not a forecast of the next repo spike; what it provides is the term the balance-sheet debate has lacked, the compensation that holders of claims on the policy rate require for the distance between reserve supply and scarcity, and how that compensation moved as the balance sheet contracted, expanded, and contracted again.

Related literature and contribution. This paper connects three literatures. The first studies how the central bank's balance sheet and the supply of reserves transmit to money-market rates. The reserve-demand strand estimates a nonlinear and shifting demand for reserves and dates the approach to scarcity (Afonso et al., 2025; López-Salido and Vissing-Jørgensen, 2023; Copeland et al., 2025), and a related strand shows that reserve and collateral quantities are impounded in repo and bill spreads and in the dispersion of overnight rates, most visibly in September 2019.[4] These studies price a contemporaneous

[4]See Afonso et al. (2020), Anbil et al. (2021), d'Avernas and Vandeweyer (2024), He et al. (2022), and Fleckenstein and Longstaff (2024) on the pricing of reserve and collateral quantities in overnight markets,

spread or date the onset of scarcity. We carry the reserve-quantity link from the spot market to the term structure: reserve adequacy is compensated in claims on the policy rate over the coming quarters, and the steepening of reserve demand that this literature documents is the reason.

The second literature asks whether money-market and bond futures carry forecastable excess returns, and what drives them. Time-varying risk premia in the term structure are well documented (Campbell and Shiller, 1991; Cochrane and Piazzesi, 2005; Ludvigson and Ng, 2009; Cieslak and Povala, 2015), and futures on policy rates have long been read as risk-adjusted forecasts of Fed policy (Piazzesi and Swanson, 2008; Sack, 2004; Hamilton and Okimoto, 2011; Cieślak, 2018; Schmeling et al., 2022). The predictors in this literature are yields, macro aggregates, or belief proxies; the closest to ours are balance-sheet quantities, as when Haddad and Sraer (2020) find that banks' interest-rate exposure forecasts Treasury excess returns or Greenwood et al. (2023) tie the supply of government debt to term premia. Our predictor is an aggregate quantity set by the central bank's own operations rather than an intermediary exposure, the asset is a claim on the overnight rate rather than on duration, and the forecasting content is established out of sample as well as in.

The third literature is dynamic term structure modeling (Duffie and Kan, 1996; Dai and Singleton, 2000; Duffee, 2002; Kim and Wright, 2005; Joslin et al., 2011; Adrian et al., 2013) with unspanned factors (Ang and Piazzesi, 2003; Joslin et al., 2014; Heath, 2019; Bauer and Rudebusch, 2017). The factors it has found unspanned are macroeconomic: real activity, inflation, return-forecasting cycles. We introduce a monetary-policy quantity as the unspanned factor and give it an institutional reason to be unspanned: under the floor system, administered rates are set to hold the overnight rate away from routine movements in reserve quantities. The spanning tests establish the property, the model imposes it, and the comparison with a spanned specification in the normalization of Joslin et al. (2013) shows that the premium does not depend on it.

and Krishnamurthy and Vissing-Jorgensen (2011), Hanson and Stein (2015), and Greenwood et al. (2016) on the transmission of asset purchases to longer-term rates.

The rest of the paper proceeds as follows. Section 2 presents the data, the spanning tests, and the forecasting evidence in and out of sample. Section 3 develops and estimates the affine model, and Section 4 reports the premium it implies. Section 5 turns to the overnight market for the mechanism, and Section 6 concludes.

2 The Reserve Adequacy Factor and Futures Excess Returns

This section introduces the futures data and the reserve-adequacy factor, then establishes the two halves of the unspanned claim in turn: the curve does not span reserve adequacy, and reserve adequacy nonetheless forecasts how the curve evolves one week ahead. A horizon profile of excess returns, estimated in sample and then forecast out of sample, places the compensation at the multi-quarter horizons on which balance-sheet policy operates.

2.1 Federal Funds Futures and the Curve Factors

We build the money-market curve from CME 30-Day Federal Funds futures, each of which settles on the arithmetic average of the effective federal funds rate (EFFR) over its delivery month. From the contract-level record we assemble a weekly panel of the first twelve monthly maturities, so that the cross-section spans one to twelve months ahead. Because a funds-futures rate is a direct claim on the Fed's operating target, this maturity window isolates the front of the term structure, where reserve conditions operate, while remaining free of the duration risk, safe-haven demand, and long-horizon growth expectations that dominate longer Treasury yields.[5]

Each settlement price is quoted as $100 - R$, where R is the implied rate in percentage points; we work with the implied rate directly. The ample-reserves regime, 2010–2025,

[5]Contract-level settlement prices, volumes, and open interest are drawn from CME historical files. We validate each contract against four gates (level, cross-sectional discrimination, argmin location, and coverage) and exclude the current-delivery contract from the panel, whose rate is mechanically pinned by the realized EFFR already accrued within the month. Construction details are in Appendix A.

supplies 835 weekly observations and spans two balance-sheet runoff episodes, 2017–2019 and 2022–2025, within a single contraction, expansion, and contraction. We use two samples for distinct purposes: the full panel for summary statistics and the description of the state, and the balanced twelve-contract panel, the 808 weeks from 8 January 2010 to 27 June 2025 on which every maturity clears the availability and liquidity screens, for the spanning tests, the predictive regressions, and the affine model. The second runoff is still under way at the end of the balanced panel, so we do not treat it as a completed cycle. The scarce-reserve regime of the 2000s is excluded: before 2008 reserve balances were a small fraction of a percent of GDP, and their weekly variation was operational noise around a trend rather than the slow-moving adequacy signal the multi-quarter horizons require.

The cross-sectional structure of the ample-sample curve is nearly one-dimensional. The first principal component of the twelve-maturity panel accounts for about 99 percent of its variation and loads almost uniformly across maturities, a level factor tracking the expected policy path, with a slope-like second component and a small third component carrying the remainder (Table 1). The near-degeneracy is itself informative: the overnight-rate curve is, to first order, a single expected-policy-path factor, which leaves reserve adequacy as a candidate second dimension the curve does not accommodate. For pricing we retain three factors, the minimum required to price the twelve maturities to about two basis points (Section 3); for the spanning and predictive regressions that follow, the leading components are the natural controls.

2.2 Measuring Reserve Adequacy

We measure reserve adequacy as aggregate reserve balances scaled by nominal GDP,

$$\mathrm{RA}_t \;=\; \frac{\text{Total Reserve Balances}_t}{\text{Nominal GDP}_t}, \tag{1}$$

the yardstick Federal Reserve officials use to discuss the ample-reserves boundary. A nominal-dollar measure would conflate the policy stance with secular growth, since nominal GDP more than doubled over the sample, so a fixed dollar level of reserves would mechanically look less adequate over time. Governor Waller places the boundary below which reserves would cease to be ample near 9 percent of GDP, and attributes the September 2019 stress to reserves having fallen below 8 percent (Waller, 2025); the New York Fed uses a comparable 8-to-10 percent range (Clouse et al., 2025). Scaling by GDP is also grounded in the mechanism: a larger economy generates higher payment volumes and greater collateral-transformation needs, raising the banking system's structural demand for settlement balances (Afonso et al., 2025; Copeland et al., 2025).[6]

Because the factor must be usable in real time, we construct RA_t from data available at each date t: weekly reserve balances enter with their publication lag, and the GDP denominator is the latest released value, taken from the real-time (ALFRED) vintage at its actual release date, rather than an interpolated or revised path. This vintage discipline matters for everything that follows: a look-ahead denominator would smooth exactly the variation we later ask the factor to predict with.

Figure 1 plots the reserves-to-GDP ratio and the funds-futures curve over the ample sample. The ratio ranges from a low of 6.5 percent in September 2019, the week of the repo spike, to a peak of 18.8 percent at the height of pandemic QE, and stands at 9.5 percent at the end of the sample, just above Waller's boundary. Three phases are visible: the balance-sheet runoff of 2017–2019 that carried reserves down toward the ample floor, the pandemic-era expansion of 2020–2021, and the renewed drawdown from mid-2022 as the second round of tightening began. For use as a state variable we take the logarithm of the ratio and demean it within the estimation sample; we write m_t for this factor and use

[6]Reserve balances are not the only overnight claim on the Fed that the private sector holds. Money-market funds and other counterparties place cash in the overnight reverse repurchase facility, and that balance is drawn down before reserves are when the balance sheet shrinks. Appendix E.1 reports the spanning, one-step and excess-return results for the broader measure that adds overnight reverse repurchase balances to the numerator; each is at least as strong. We keep reserves over GDP as the headline measure because it is the quantity the policy debate is conducted in.

it throughout. One standard deviation of m_t is 0.25 log points, and coefficients on it are reported in standardized units in the reduced-form regressions that follow. Table 1 reports summary statistics for the contract series and for RA.

[Insert Figure 1 here]

[Insert Table 1 here]

2.3 Spanning

Following the unspanned-factor framework of Joslin et al. (2014), we first ask whether reserve adequacy is recoverable from the contemporaneous curve. We project RA_t on the leading principal components of the funds-futures panel,

$$\mathrm{RA}_t \;=\; \gamma_0 + \sum_{k=1}^{K} \gamma_k \, \mathrm{PC}_{k,t} + \mathrm{RA}_t^{\perp}, \tag{2}$$

and examine how much of RA the curve explains. In levels the curve accounts for little, an R^2 of 0.07 with three components, and in innovations, the relevant object once the common near-unit-root trend is removed, it accounts for essentially nothing: a spanning R^2 of about 0.01 that barely moves as components are added (Table 2). Roughly ninety-nine percent of the innovation variation in reserve adequacy is invisible to the cross-section of futures prices. In the terminology of Joslin et al. (2014), RA is unspanned, which is the first half of the definition. The flatness of the R^2 as factors accumulate is the signature Bauer and Rudebusch (2017) associate with a genuinely unspanned factor: adding curve dimensions does not begin to recover it. The result does not depend on the scaling of reserves; the ratios to bank assets and to deposits, the broader measure that includes overnight reverse repurchase balances, and log reserves are all unspanned to the same degree (Appendix Table E.2). The floor system supplies the institutional reason to expect this, since administered rates are set to hold spot funding rates away from routine movements in reserve supply. We treat that as motivation rather than proof: the affine model of

Section 3 imposes a separate risk-neutral exclusion, which this finding motivates but does not imply, and Section 3.6 shows that relaxing the exclusion and pricing reserve adequacy as a spanned factor gives it a curve loading of essentially zero.

The orthogonal residual $\mathrm{RA}_t^{\perp}$ inherits the near-unit-root persistence of reserves, with a weekly first-order autocorrelation of 0.99. This persistence is what allows a one-week forecasting relation to accumulate into a multi-quarter premium, and it is the property the model of Section 3 is built around.

[Insert Table 2 here]

2.4 Reserve Adequacy in the Dynamics of the Curve

Unspannedness alone does not make reserve adequacy interesting. A state variable that is invisible to today's cross-section matters for pricing only if it moves the curve's subsequent evolution; if it does not, it can be dropped from the model without loss. Joslin et al. (2014) make this the substantive half of the unspanned-factor definition, and Heath (2019) tests it directly in a futures market by asking whether a macroeconomic state forecasts next-period changes in the curve over and above the curve itself.

Table 3 runs that regression here. The dependent variable is the one-week change in a leading principal component of the funds-futures curve; the controls are the contemporaneous curve, entered three ways, as two components, three components, and all twelve contract rates; and the coefficient of interest is on reserve adequacy. Because the horizon is a single period, the forecast windows do not overlap.

Reserve adequacy forecasts the curve. A one-standard-deviation increase raises the expected one-week change in the level component by about 0.03 principal-component units with a t-statistic of up to 3.4, and it raises the slope component as well, significantly at the 10 percent level in every column and at the 5 percent level in five of the six. Neither result depends on how the curve is controlled for: entering all twelve contract rates rather than two principal components leaves the coefficients essentially unchanged. The third

principal component is not forecastable and we omit it.

The even-numbered columns add the variables this literature uses to forecast funds-futures excess returns, payroll growth, the credit spread, and volatility (Piazzesi and Swanson, 2008; Cieślak, 2018). On the level factor the reserve-adequacy coefficient barely moves and none of the three predictors enters significantly; what widens is the standard error. On the slope factor the coefficient attenuates somewhat, and payroll growth enters with the sign that literature reports. Reserve adequacy is therefore not a repackaging of the conditioning information already known to forecast this market. This is the property the affine model of Section 3 formalizes: there the same relation is the block $\mathbf{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}}$ of the physical dynamics, and Section 4 tests whether it is zero.

[Insert Table 3 here]

2.5 The Horizon Profile of Excess Returns, In and Out of Sample

The one-period result says reserve adequacy moves the curve. It does not say over what horizon the compensation for that movement accumulates, and the horizon is where the economics is: balance-sheet policy operates over quarters, not weeks. Following Piazzesi and Swanson (2008), the h-quarter excess return is the futures-implied rate minus the overnight rate subsequently realized over the contract's accrual window,

$$rx_t^{(h)} \;=\; f_t^{(h)} - \bar{r}^{\,\text{realized}}_{t+(h-1)\mathrm{Q}\to t+h\mathrm{Q}}, \qquad h = 1,\ldots,5, \tag{3}$$

so that a positive rx is compensation earned by a long position and a negative loading on RA means abundant reserves map to lower subsequent excess returns. Adding the spanning residual to the curve factors,

$$rx_t^{(h)} \;=\; \beta_0 + \sum_{k=1}^{K} \beta_k \,\mathrm{PC}_{k,t} + \beta_{\mathrm{RA}}\,\mathrm{RA}_t^{\perp} + \varepsilon_t^{(h)}, \tag{4}$$

gives the profile in Panel A of Table 5. At one quarter the incremental R^2 is smallest, which is the institutionally expected outcome: the front of the curve is pinned by administered rates and monitored by the Desk's realized gauges, so whatever scarcity information exists there is already in the dashboard. From two quarters outward it rises monotonically, reaching 0.38 at five quarters, and the loading is negative throughout and grows in magnitude with the horizon, from about 6 basis points per standard deviation of reserve adequacy at one quarter to 72 at five. Hamilton and Okimoto (2011) document the same horizon concentration in fed funds futures returns of an earlier era. These windows overlap by up to sixty-five weeks and the regressor is near-integrated, so we read the in-sample profile for its shape and sign; Section 4 recovers the same profile from the model, where it follows from the persistence of reserve adequacy rather than from an overlapping regression.

Table 4 states the four-quarter relation the way the return literature states it (Piazzesi and Swanson, 2008; Hamilton and Okimoto, 2011): the excess return on the funds-futures ladder regressed on the curve through the model's own pricing-state coordinates, on reserve adequacy, and on the same three return predictors. A one-standard-deviation decline in reserve adequacy is worth about half a percentage point of four-quarter excess return, and the loading moves from -0.53 to -0.50 when the predictors are added; the result is robust to realistic release lags on the macro series.

[Insert Table 4 here]

An in-sample profile can be a fitted artifact of a persistent regressor. The forecasting exercise of Joslin et al. (2014) asks the question that cannot be: does reserve adequacy improve forecasts made in real time, with parameters estimated only on data available at the forecast date? Panel B of Table 5 reports pseudo-out-of-sample forecasts of the h-quarter excess return from an expanding window, comparing a model with the two curve factors and reserve adequacy against the curve alone, using the out-of-sample R^2 of Campbell and Thompson (2008) and the test of Clark and West (2007) for nested models.

Reserve adequacy improves the forecasts, and by a wide margin at the horizons where the premium lives. At two quarters the out-of-sample R^2 relative to the curve-only benchmark is 25 percent, and at four quarters it is 35 percent, with Clark–West statistics that are significant at conventional levels at every horizon. The standard return predictors, added to the curve on their own, do not improve the forecasts at any horizon, and adding reserve adequacy on top of them produces gains of the same size as adding it to the curve alone. The gain is not the product of a single episode: when the forecasts made in 2023 and 2024, the years in which reserves fell back toward the ample boundary, are set aside, the out-of-sample R^2 at two and four quarters is still 11 and 16 percent, and the Clark–West statistics are larger, not smaller.[7]

[Insert Table 5 here]

Both halves of the unspanned-factor condition are now in place, and the second holds out of sample. Reserve adequacy is not recoverable from the contemporaneous cross-section, and it forecasts that cross-section's subsequent evolution and the returns that follow from it. Joslin et al. (2014) show that a factor with these two properties can be excluded from the pricing map and still shape the compensation embedded in longer-dated claims. Section 3 builds that model and measures the compensation directly.

3 An Affine Model with Unspanned Reserve Adequacy

Section 2 establishes that reserve adequacy is unspanned by the curve and forecasts its evolution. What a reduced form cannot deliver is the size of the compensation attached to that forecast, since an overlapping average of realized returns measures a premium only through the noise of the returns themselves. This section builds the pricing model that

[7] Appendix Table E.1 reports the same comparison with a rolling five-year window, with forecasts beginning in 2017, within the 2015–2019 and 2020–2025 subperiods, with three curve factors, and with the principal-component loadings estimated only from prices available at each forecast date rather than on the full sample; Appendix Figure E.1 plots the cumulative difference in squared forecast errors in the manner of Welch and Goyal (2008). The gain accrues slowly through the first runoff and mostly through the second.

asks the question directly: what compensation an arbitrage-free curve must be carrying at each date, and how much of that compensation moves with reserve conditions. The model is a discrete-time affine term structure model for federal funds futures in which reserve adequacy enters as an unspanned state variable in the sense of Joslin et al. (2014). It follows the canonical Gaussian framework of Dai and Singleton (2000) and Duffee (2002) under the normalization of Joslin et al. (2011), adapted to the settlement conventions of monthly-average contracts and estimated in the two-step manner of Joslin et al. (2014) and Heath (2019). Throughout, RA_t denotes the reserve-adequacy factor m_t of Section 2.2, the logarithm of the reserves-to-GDP ratio, demeaned within the estimation sample, one standard deviation of which is 0.25 log points.

3.1 State Space

Time is discrete and indexed by t, with one period equal to one business day, the frequency at which the underlying rate is set. Let $X_t \in \mathbb{R}^3$ be the latent pricing state and let

$$Z_t = \begin{pmatrix} X_t \\ \mathrm{RA}_t \end{pmatrix} \tag{5}$$

be the full state at date t. The three pricing factors are the curve factors of Section 2.1: we fix a loading matrix $\boldsymbol{W}$ once and for all as the first three principal components of the twelve-contract ladder, so that $P_t = \boldsymbol{W} f_t$ is a fixed set of portfolios of futures rates, and X_t is the affine rotation of P_t implied by the pricing map below.[8] Figure 2 plots the three loadings: a near-uniform level component, a monotone slope, and a hump-shaped curvature. Reserve adequacy is the fourth state variable and, as the two exclusion restrictions below make precise, it is not a fourth pricing factor.

[8] $\boldsymbol{W}$ is computed on the per-contract demeaned covariance matrix of the ample sample and rescaled row-wise as in Joslin et al. (2014); it is frozen before estimation and is never re-extracted. Because $\boldsymbol{W}$ defines the factors rather than estimating them, this is the standard convention in the unspanned-factor literature.

[Insert Figure 2 here]

Under the physical measure $\mathbb{P}$ the state follows a first-order vector autoregression,

$$Z_{t+1} \;=\; \boldsymbol{\mu}^{\mathbb{P}} + \boldsymbol{\Phi}^{\mathbb{P}} Z_t + \boldsymbol{\Sigma}_Z \, \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim \mathcal{N}(\mathbf{0}, I_4), \tag{6}$$

with $\boldsymbol{\Phi}^{\mathbb{P}}$ left unrestricted. The unrestricted form is not a matter of generality for its own sake: the entire unspanned mechanism operates through the off-diagonal block $\boldsymbol{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}}$, the effect of today's reserve conditions on the future curve factors. If that block is zero the model collapses to a curve-only term structure model in which reserve adequacy is irrelevant to prices; if it is not, reserve adequacy moves the physical forecast of future short rates without moving today's prices, and the wedge is a term premium.

The short rate is the sum of the pricing factors,

$$r_t \;=\; \boldsymbol{\iota}' X_t, \qquad \boldsymbol{\iota} = (1,1,1)', \tag{7}$$

which is the Joslin et al. (2011) normalization rather than a substantive restriction: the intercept and scale of the latent state are absorbed into the risk-neutral parameters below. What *is* substantive is that RA_t does not appear in (7). The exclusion has an institutional reading specific to the floor system. The Federal Reserve sets administered rates—interest on reserve balances and the overnight reverse repurchase rate—precisely so that the effective funds rate is determined within a corridor rather than by the quantity of reserves; as long as reserves remain ample, the overnight rate is where the administered rates put it, not where the supply of balances puts it. Reserve adequacy is therefore excluded from the contemporaneous short rate by policy design, and it reaches future short rates only through the physical dynamics (6).

Under the risk-neutral measure $\mathbb{Q}$ the pricing factors evolve as

$$X_{t+1} \;=\; \boldsymbol{\mu}^{\mathbb{Q}} + \boldsymbol{G}\, X_t + \boldsymbol{u}^{\mathbb{Q}}_{t+1}, \qquad \boldsymbol{\mu}^{\mathbb{Q}} = (\kappa, 0, 0)', \quad \boldsymbol{G} = \mathrm{diag}(g_1, g_2, g_3), \tag{8}$$

with $g_1 > g_2 > g_3$. Two features of (8) carry the model's identification. First, the risk-neutral feedback matrix is diagonal and the risk-neutral intercept $\boldsymbol{\mu}^{\mathbb{Q}}$ is a single scalar κ in the first coordinate and zero elsewhere: this is the Joslin et al. (2011) canonical form, in which $(\boldsymbol{G}, \kappa)$ exhaust the $\mathbb{Q}$-block. Second—and this is the unspanned restriction proper—RA_t has no column in (8). Futures prices are functions of X_t alone. Together with the exclusion from (7), this makes reserve adequacy unspanned in the model in exactly the sense in which Section 2.3 found it unspanned in the data: it is not recoverable from the cross-section, because the cross-section does not depend on it.

The change of measure follows Duffee (2002), with the price of risk affine in the state,

$$\lambda_t \;=\; \boldsymbol{\Lambda}_0 + \boldsymbol{\Lambda}_1 Z_t, \qquad \boldsymbol{\Lambda}_1 \;=\; \boldsymbol{\Sigma}_X^{-1}\left(\boldsymbol{\Phi}^{\mathbb{P}}_{X,\cdot} - \boldsymbol{\Phi}^{\mathbb{Q},\text{ext}}\right), \qquad \boldsymbol{\Phi}^{\mathbb{Q},\text{ext}} = \begin{pmatrix} \boldsymbol{G} & \mathbf{0} \end{pmatrix}, \tag{9}$$

where $\boldsymbol{\Phi}^{\mathbb{Q},\text{ext}}$ extends the risk-neutral feedback of the pricing factors to the full state by appending a zero column for RA, and $\boldsymbol{\Sigma}_X$ is the pricing-factor block of $\boldsymbol{\Sigma}_Z$. The reserve-adequacy column of $\boldsymbol{\Lambda}_1$ is therefore proportional to $\boldsymbol{\Phi}^{\mathbb{P}}_{X,\text{RA}}$: because RA enters the physical dynamics but is absent from the risk-neutral ones by construction, the entire gap in that column is physical, uncontaminated by any risk-neutral parameter. Whether its individual entries are separately estimable is a different question, and Section 3.6 answers it in the negative; what the paper uses are quantities defined on the column as a whole. The reserve-adequacy *row* of $\boldsymbol{\Lambda}_1$—how the compensation for reserve shocks itself varies—is not identified, since RA's own risk-neutral dynamics never enter a price; we do not estimate or report it. This asymmetry is the standard price of the unspanned specification (Joslin et al., 2014; Heath, 2019) and it is worth stating plainly what the identified column does and does not mean: it measures how reserve conditions shift the compensation embedded in the curve for bearing the priced curve risks. The compensation it measures attaches to the state of reserve adequacy through the curve risks the model prices; the price of risk on reserve shocks themselves is not identified in an unspanned model and is not reported.

3.2 Pricing Federal Funds Futures

A 30-day federal funds futures contract for delivery month m settles on the arithmetic average of the effective funds rate over the *calendar* days of the month. Because the settlement rate is affine in the underlying, the futures rate is exactly the risk-neutral expectation of that average, with no convexity correction, a simplification relative to compounded-average contracts, whose pricing requires one; the covariance of the Gaussian innovations $\boldsymbol{u}_{t+1}^{\mathbb{Q}}$ in (8) therefore never enters a price.

The averaging convention requires care at month boundaries. The effective funds rate is published for business days only, and a weekend or holiday carries the rate of the preceding business day, which may fall in the previous calendar month. Let N_m be the number of calendar days in month m. Seen from a quotation date t, let $\mathcal{H}_m$ be the set of horizons h such that the rate set on day $t+h$ is counted in month m, and let the weight $\omega_{m,h}$ be the number of calendar days of month m whose rate is set on day $t+h$, divided by N_m; both depend on t, which we suppress. A business day followed by a weekend therefore carries three days' weight rather than one, and when the month opens on a weekend or holiday part of its weight belongs to the last business day of the previous month. The weights satisfy $\sum_{h\in\mathcal{H}_m} \omega_{m,h} = 1$, and the settlement rate is $\sum_{h\in\mathcal{H}_m} \omega_{m,h}\, r_{t+h}$, which we use in place of uniform weights on the business days within the month.[9]

Proposition 1 (Funds Futures Pricing)**.** *Under the short rate* (7) *and the risk-neutral dynamics* (8)*, the rate implied by the futures contract for delivery month m is affine in the pricing state,*

$$f_t^{(m)} \;=\; \bar{A}_t^{(m)} + \big(\bar{B}_t^{(m)}\big)' X_t, \tag{10}$$

[9]The difference between the two weightings is immaterial on average and material where it matters: the discretization error of a within-month approximation has a median of about 0.1 basis points but reaches 13 basis points at month boundaries, which is an order of magnitude larger than the pricing errors the model is asked to explain. All business-day counts are anchored to the quotation date rather than to the weekly date stamp, and the business-day calendar is the Federal Reserve's.

with loadings

$$\left(\bar{B}_t^{(m)}\right)' = \sum_{h \in \mathcal{H}_m} \omega_{m,h}\, \boldsymbol{\iota}' \boldsymbol{G}^h, \tag{11}$$

$$\bar{A}_t^{(m)} = \kappa \sum_{h \in \mathcal{H}_m} \omega_{m,h}\, S_h, \qquad S_h = \sum_{j=0}^{h-1} g_1^j. \tag{12}$$

The proof is in Appendix B. Three implications organize what follows. First, the loadings are forward averages of powers of a diagonal matrix, so each factor's contribution to a futures rate decays at its own rate and is averaged over the delivery month; this forward-average structure is what makes principal components of the panel a consistent estimator of the factor space (Joslin et al., 2011; Bai and Ng, 2002) and justifies fixing $\boldsymbol{W}$ ahead of estimation. Second, the loadings are date-specific: the horizons $h \in \mathcal{H}_m$ to the days of the delivery month shorten as the quotation date moves through the month, so $\bar{A}_t^{(m)}$ and $\bar{B}_t^{(m)}$ are recomputed at every date rather than indexed by maturity alone. Third, and most importantly, RA_t appears nowhere in (10). Whatever reserve adequacy does in this model, it does not do it to today's prices.

3.3 The Reserve-Adequacy Component of the Term Premium

The term premium is the difference between the futures rate and the physical expectation of the settlement average it will pay,

$$\mathrm{TP}_t^{(m)} = f_t^{(m)} - \mathbb{E}_t^{\mathbb{P}}\left[\sum_{h \in \mathcal{H}_m} \omega_{m,h}\, r_{t+h}\right], \tag{13}$$

where the first term is the risk-neutral expectation of the same object by no-arbitrage. The sign convention matches Section 2.5: $\mathrm{TP}_t^{(m)}$ is the expected excess return $\mathbb{E}_t[rx]$ on a long position in the contract, so a negative loading on RA means that abundant reserves go with a low premium.

The two expectations in (13) are taken with respect to different information. The risk-neutral one runs on X_t alone, through (8). The physical one runs on the full state Z_t, by iterating the VAR (6) to the horizon of each day in the delivery month, with the weights $\omega_{m,h}$ of Proposition 1 carried through unchanged, so that the two legs are averaged over exactly the same days. Substituting both, the term premium is affine in the state,

$$\mathrm{TP}_t^{(m)} \;=\; \beta_0^{(m)} + \big(\beta_X^{(m)}\big)' X_t + \beta_{\mathrm{RA}}^{(m)} \mathrm{RA}_t, \tag{14}$$

and the reserve-adequacy loading has an explicit form,

$$\beta_{\mathrm{RA}}^{(m)} \;=\; -\sum_{h \in \mathcal{H}_m} \omega_{m,h}\, \boldsymbol{\iota}' \Big[\big(\boldsymbol{\Phi}^{\mathbb{P}}\big)^h \Big]_{X,\mathrm{RA}}, \tag{15}$$

where, for a matrix on the full state, $[\,\cdot\,]_{X,\mathrm{RA}}$ denotes its reserve-adequacy column restricted to the rows of the three pricing factors, so that $\beta_{\mathrm{RA}}^{(m)}$ is minus the cumulative effect of a reserve-adequacy shock on the expected path of the short rate over the delivery month. Equation (15) is the structural counterpart of the horizon profile in Section 2.5. A one-period feedback coefficient from RA to the curve factors may be small; what matters for a contract m months out is its accumulation over h periods, and because reserve adequacy is close to a unit root that accumulation compounds rather than decays, building more than proportionately with the horizon. The model therefore predicts a premium component that is negligible at the front of the curve and builds with maturity. The estimation does not fit that pattern. It follows arithmetically from a persistent state variable that enters the physical forecast and not the price.

Because (14) is exact rather than a projection, the contribution of reserve adequacy to the term premium can be attributed directly. We define the reserve-adequacy attribution

for the twelve-month term premium as

$$\text{share}_{\text{RA}} = \frac{\beta_{\text{RA}}^{(12)} \operatorname{Cov}\left(\text{RA}_t,\ \text{TP}_t^{(12)}\right)}{\operatorname{Var}\left(\text{TP}_t^{(12)}\right)}, \tag{16}$$

the fraction of the model-implied variance of the twelve-month premium attributable to reserve adequacy. This is a *covariance* attribution rather than an orthogonal variance decomposition: the state variables are correlated, so the curve and reserve-adequacy contributions need not sum to one, and in finite samples the attribution can fall outside $[0, 1]$. This scalar, and the test of whether $\boldsymbol{\Phi}^{\mathbb{P}}_{X,\text{RA}} = 0$ that underlies it, are the model's two answers to the question Section 2 left open; both are reported in Section 4.

3.4 Estimation

We estimate the model in two steps, risk-neutral parameters first, on the 808 weeks of the ample regime for which the complete twelve-contract ladder and RA are simultaneously available, 8 January 2010 through 27 June 2025. The first step fits the risk-neutral block to the cross-section: each of the twelve contract series carries an independent Gaussian pricing error, the three portfolios $P_t = \boldsymbol{W} f_t$ are priced exactly in the observable-factor normalization of Joslin et al. (2011), and (g_1, g_2, g_3, κ) are estimated by concentrated quasi-maximum likelihood. The second step takes those estimates as given, inverts the pricing state from the cross-section date by date, and estimates the physical VAR (6) by ordinary least squares on $\hat{Z}_t = (\hat{X}_t', \text{RA}_t)'$. The two steps use the data at different frequencies: the risk-neutral block is fitted at the business-day step of the model, while the physical VAR is estimated on the weekly observations, so that the estimated $\boldsymbol{\Phi}^{\mathbb{P}}$ is a one-week transition and a horizon of h business days in (15) is iterated over $w(h) = \text{round}(h/5)$ weekly steps (Appendix B). Appendix D gives the estimator in full, together with the conventions it declares, the treatment of the current-delivery contract, and the bias correction described

next.

One estimation choice belongs in the text, because it moves a headline magnitude. Least squares understates the persistence of a near-unit-root state in a sample of this length, and the understatement propagates into (15), where it compounds over horizons. The least-squares estimates are the ones the text carries, as in Joslin et al. (2014) and Heath (2019); Table 7 reports the bias-corrected attribution of Bauer et al. (2012) beside them for comparison.

3.5 Estimation Results

Three factors price the twelve-maturity ladder to a root-mean-square error of 1.6 basis points on average and no more than 2.27 at any maturity, a tight fit for a market whose quoted rates move in half-basis-point increments, and one obtained with a risk-neutral block of four parameters. Two factors do not suffice: they price the interior of the ladder adequately but miss both ends by about six and a half basis points, a failure that is a property of the curve rather than of the estimator (Appendix D). Panel A of Table 6 reports the risk-neutral estimates, which Appendix D discusses.

[Insert Figure 3 here]

[Insert Table 6 here]

What matters economically is the physical block. The system is highly persistent, and reserve adequacy is its most persistent element: a weekly autoregressive coefficient of 0.9939, a half-life of about two years. This is the persistence that the overlapping regressions of Section 2.5 inherit. In the model it is not a nuisance but the mechanism. A state variable this persistent, entering the physical forecast of an object that is priced without it, generates a premium component that accumulates over horizons rather than dissipating, which is why the reserve-adequacy loading of Section 3.3 builds with the contract horizon instead of washing out.

The block that carries the mechanism is the reserve-adequacy column of $\mathbf{\Phi}^{\mathbb{P}}_{X,RA}$, reported in Panel B of Table 6. Its effect on the level-like pricing factor, the one coordinate separated from the near-degenerate pair, is 0.041; the column's effect on the expected short rate is 0.0062, a quantity invariant to how that pair is parameterized. Whether that column is zero, and how large a term premium it generates, are the questions Section 4 takes up.

3.6 Identification and Specification Checks

The model is estimated on a single balance-sheet cycle with a near-unit-root state. Three features of the estimates govern how its output should be read. Appendix D documents each in full.

Reserve adequacy is unspanned inside the model as well as outside it. Relaxing both exclusions and pricing reserve adequacy as a spanned factor, an exact affine function of four latent factors with free coefficients in the normalization of Joslin et al. (2013), assigns it a curve loading of at most 0.15 basis points per standard deviation and leaves the twelve-month premium essentially unchanged, with a correlation of 0.999 between the two models' series (Appendix Table E.3). The likelihood gain from pricing it is no larger than the gain from pricing a placebo series that carries no reserve information, so the exclusion is a restriction the cross-section is content to bear.

The largest risk-neutral eigenvalue exceeds unity by about three standard errors, a familiar consequence of estimating a Gaussian affine model over a sample half of which sits at the zero lower bound, where the level factor has no mean reversion to reveal. We therefore report risk-neutral quantities only within the twelve-month ladder, and no unconditional or long-run quantity at all; on the post-2016 subsample, which excludes the first zero-lower-bound episode, the root falls below one.[10]

[10]The estimation window spans two zero-lower-bound episodes (2010–2015 and 2020–2021), over which a Gaussian affine model permits negative short rates. Appendix C reports an illustrative shadow-rate counterfactual: holding the estimated dynamics fixed and imposing $r_t = \max(\underline{r}, \boldsymbol{\iota}' X_t)$, the model-implied reserve-adequacy contribution to the term premium is unchanged when the funds rate is above 2.5 percent, a standard-deviation ratio of 1.00, is about nine-tenths of its affine size between 0.25 and 2.5 percent, and

The second and third risk-neutral eigenvalues are not separately identified: the curve does not distinguish a slope factor from a curvature factor over this sample. We therefore report no quantity defined coordinate by coordinate within that pair: no individual risk-price loading, and no split of the premium into slope and curvature contributions. The aggregates the paper does report are invariant to how the pair is parameterized (Appendix D).

4 The Reserve-Adequacy Term Premium

Section 3 left the model with two objects to report. One is the reserve-adequacy share of the twelve-month term-premium variance, equation (16). The other is the restriction that generates it, $\mathbf{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0$: the hypothesis that reserve conditions do not enter the physical forecast of the curve at all, in which case the reserve-adequacy loading of the premium is identically zero. This section reports both, traces the loading across horizons and the resulting premium across the balance-sheet cycle, and closes with what the evidence does and does not establish.

4.1 The Size of the Component

Reserve adequacy accounts for about four-fifths of the model-implied variance of the twelve-month term premium. The covariance attribution is 0.79 (Table 7).[11] The restriction $\mathbf{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0$ is rejected at the 5 percent level: its likelihood-ratio statistic is 11.0 on three degrees of freedom. A contribution of this size is the paper's central quantitative result.[12]

compresses to about six-tenths at the bound, where a floored short rate mechanically removes the rate risk the premium compensates. The reserve-adequacy share of the term premium is 0.79 in the affine model and 0.69 under a ten-basis-point floor.

[11] Under the bias-corrected physical dynamics of Bauer et al. (2012) the attribution is 0.53 (Table 7); Section 3.4 explains why the least-squares estimate is the one carried in the text.

[12] The two-factor version of the model in Appendix D rejects the same restriction more sharply, on two degrees of freedom. We report the three-factor result because a premium measured in a model that misprices the ends of the ladder would not be credible, not because the restriction is harder to reject there.

[Insert Table 7 here]

4.2 The Horizon Profile

The reserve-adequacy loading of the premium builds with the contract horizon. In the excess-return convention a one-log-point decline in reserve adequacy raises the premium by about 18 basis points at one quarter and by 160 at four (Table 8), and the standard deviation of the component it generates rises from 4.6 to 40.6 basis points over the same horizons. This is the structural counterpart of the reduced-form profile of Section 2.5, and it arises for the reason equation (15) makes plain: a persistent state variable that enters the physical forecast but not the price contributes a premium that accumulates as the forecast lengthens.

The same table places the model loading beside the reduced-form predictive slope of Section 2.5, in the same convention. The two agree in sign and grow together across horizons, the model loading running at about four-fifths of the reduced-form slope at every horizon. We show this as a consistency check on the structural estimate rather than as a second measurement; the standard errors of the reduced-form slopes are those of Table 5.

[Insert Table 8 here]

4.3 The Premium across the Balance-Sheet Cycle

Evaluating the reserve-adequacy component along the realized path of the state gives a series we write $\mathrm{ERP}_t^{\mathrm{RA}}$ and plot in Panel A of Figure 4. Because reserve adequacy is the only unspanned state variable and enters the premium linearly, this series is by construction a rescaling of reserve adequacy itself, with a correlation of -1 with m_t. We read it descriptively, as a way of putting the reserve series on the axis of the premium it loads.

So read, it traces the balance-sheet cycle (Table 9). It is near zero through the abundant

early 2010s, widens to an average of 36 basis points over the 2017–2019 runoff as reserves fell toward the ample boundary, turns sharply negative in the pandemic expansion of 2020–2021 as reserves flooded in, and sits mildly negative through the second round of tightening. Over the full sample it is widest in the weeks around the September 2019 repo spike and narrowest in the reserve abundance of September 2021, and it co-moves with the spread of the funds rate over the administered floor, the scarcity gauge the Desk watches. The out-of-sample gains of Section 2.5 have the same timing: they accrue through the two runoffs, when reserves were falling toward the boundary and the component was widening.

The economics is a scarcity premium. In the excess-return convention the correlation between reserve adequacy and the full term premium is −0.91: scarcer reserves, a higher premium. This is the object the model isolates, compensation for bearing reserve-scarcity risk over the life of the contract rather than a revision to the expected policy path (Piazzesi and Swanson, 2008; Hamilton and Okimoto, 2011), and Section 5 shows where the risk comes from.

[Insert Table 9 here]

[Insert Figure 4 here]

4.4 The Scope of the Evidence

The premium we measure is model-implied. It is the wedge between the futures rate and the physical forecast of the funds rate that an arbitrage-free affine model attributes to reserve adequacy, estimated over one balance-sheet cycle with a near-unit-root state. A survey-based alternative is not available for this market: the New York Fed's Survey of Primary Dealers reports the modal path of the policy rate rather than its mean, so the gap between the futures rate and the survey path mixes the premium with the skew of the rate distribution, and the survey's coverage since 2018 spans a single balance-sheet cycle that

moves with the rate cycle.[13]

Within the model, the component does not depend on the unspanned restriction. Pricing reserve adequacy as a spanned factor leaves the twelve-month premium essentially unchanged and gives the factor a curve loading of close to zero (Section 3.6), so the premium is generated through the physical forecast rather than through a mispricing of the current curve. And the state that generates it carries information about where funding conditions are going: in the second runoff, lower reserve adequacy forecast a tighter repo spread over the following quarter beyond what the current spread and the curve forecast (Appendix Table E.5), which is the direction a scarcity premium requires.

5 Why Reserve Adequacy Is Priced: Reserve Flows and the Overnight Rate

The model of Sections 3 and 4 measures the compensation attached to reserve adequacy. It does not say why a market in claims on the overnight rate should demand that compensation when the floor system is designed to keep reserve quantities out of the overnight rate. This section asks when that design holds and when it does not, using the overnight market itself. The answer has three parts. A given withdrawal of reserves moves the funds rate only when reserves are scarce, so the demand for reserves steepens as adequacy falls, the shape Afonso et al. (2025) estimate directly. The same withdrawals leave the futures curve untouched in both states, so reserve flows are unspanned by the curve. And what the curve responds to is the level of adequacy, which decides how much the next withdrawal will move the rate. Section 5.1 reports the slope of reserve demand by state; Section 5.2 contrasts the response of the spot rate with that of the curve and shows that the level, not

[13] Over 2018–2025 the survey rounds coincide with the zero lower bound at peak reserves, where the modal path sits at the floor while futures price a lift-off tail, and with the two runoffs at turning points where the market led dealers' modal view. The gap between the futures rate and the modal survey path therefore cannot be separated from reserve adequacy in that sample, and we do not use it. A mean path can be built only from the probability distributions the survey has collected since 2023.

the flow, forecasts the curve.

5.1 The Slope of Reserve Demand

The weekly change in reserve balances is not an exogenous shock to the funds market. The Federal Reserve adds reserves when funding spreads spike, as it did in September 2019 and March 2020, so a regression of the spread on realized reserve changes mixes the demand curve with the central bank's reaction to it. We use two ways around this. The first follows Hamilton (1997), who identified the liquidity effect from movements in the Treasury's balance at the Fed: changes in the Treasury General Account and in overnight reverse repurchase balances drain reserves one for one, and they are driven by the tax calendar, debt management, and money-fund allocation rather than by the Fed's response to funding conditions. We call their sum the autonomous drain and also report the Treasury balance on its own. The second uses the realized change in reserves but controls for the two components of the Fed's own balance sheet that move it, the change in total assets and the change in currency in circulation, so that the remaining variation in reserves is the part the autonomous factors generate.

The regression is weekly, on the Wednesday weeks of the H.4.1 release, with the change in the spread of the effective funds rate over the interest rate on reserves as the dependent variable, in basis points. Each flow is measured in percent of GDP and interacted with an indicator for reserve adequacy below ten percent of GDP, the boundary Waller (2025) places on the ample region. Standard errors are Newey–West with four lags.

Reserve flows move the overnight rate only when reserves are scarce. Table 10, Panel A, reports the slope in the two states for each of the three flows. When reserves exceed ten percent of GDP the slopes are a third of a basis point or less per percent of GDP and of mixed sign. Below ten percent an autonomous drain of one percent of GDP raises the spread by about 2.2 basis points, and the difference between the two states has a t-statistic of 3.6. The Treasury balance carries the precise part of this effect, and realized reserve

changes, once the Fed's own balance-sheet movements are controlled for, give a slope of the same size with the opposite sign, as a drain and an injection should. Figure 5 plots the slope by quintile of reserve adequacy: it is confined to the lowest quintile, where reserves ran between 6.5 and 9.5 percent of GDP, and flat across the other four. This is the steepening of reserve demand that Afonso et al. (2025) recover from a time-varying demand curve and that López-Salido and Vissing-Jørgensen (2023) build their account of quantitative tightening around, here obtained with the paper's own state variable and a shock from outside the Fed's reaction function.

[Insert Table 10 here]

[Insert Figure 5 here]

The steepening is the economics behind the loading of Section 4. In the flat region of reserve demand, a reserve drain is a non-event for the overnight rate; in the steep region the same drain moves it. Scarcity is therefore a state in which the overnight rate carries more risk per unit of reserve flow, and because adequacy is persistent, with a half-life of about two years in the estimated dynamics, the state a contract is written in is close to the state it will settle in. An investor long a twelve-month contract in the scarce state bears more rate risk over its life than one in the abundant state, and the premium of Section 4 is the compensation for that difference. The affine model does not estimate the curvature of reserve demand; it is linear in the state and transmits the state through persistence. What the overnight market supplies is the reason the state has a price.[14]

[14]The premium's loading on reserve adequacy is itself linear. Adding a kink at ten percent of GDP to the excess-return regressions of Section 2.5 adds nothing, and the loading is larger in the two runoff episodes, when reserves were falling toward the boundary (Appendix Table E.4). The nonlinearity lives in the overnight market, where reserve demand steepens; the futures curve prices the state that puts the system in the steep region.

5.2 Flows, Levels, and the Curve

If reserve flows move the overnight rate in the scarce state, do they also move the curve? Panel B of Table 10 looks at the days on which predictable drains occur: the mid-month tax dates of March, April, June, September, and December, the mid-month settlement of Treasury coupon issuance in the other months, and month-ends, the dates on which Copeland et al. (2025) document repo-rate pressure from the settlement calendar. On these days the spread of the funds rate over the administered rate rises by one to two and a half basis points more when reserves are below ten percent of GDP than when they are above, with the largest response on tax dates. The futures ladder does not respond more in the scarce state at any maturity: across the three maturity groups and the three kinds of day, no interaction with the scarce state is distinguishable from zero. A contract that settles on the average funds rate over a delivery month weeks ahead has no reason to move on a drain that reverses within days, and it does not.

What the curve responds to is the level of adequacy. Panel C returns to the one-week-ahead forecast of Section 2.4 and adds the recent change in reserve adequacy, over one week and over four, to the level. The coefficient on the level is unchanged; the four-week flow enters neither equation, and the one-week flow enters only the slope equation, at the 10 percent level. The curve prices the slow state, not the recent flow.

Together the three panels give the floor system's two faces. Day to day, administered rates do what they were designed to do: reserve flows move the overnight rate only in the scarce state, and even then the curve ignores them, because they are transitory and anticipated. Over quarters, the level of adequacy determines how often the system will sit in the steep region of reserve demand and how far a given flow will then move the rate. That is the information the futures curve carries and the premium of Section 4 compensates: a slow-moving quantity that is kept out of today's overnight rate and prices the risk in tomorrow's.

6 Conclusion

The floor system separates control of the overnight rate from the quantity of reserves, and for the overnight rate the separation holds: as long as reserves are ample, their weekly movements leave the funds rate where the administered rates put it. This paper shows that the separation does not extend to the term structure of claims on that rate. Reserve adequacy is almost entirely unspanned by the federal funds futures curve, yet it forecasts the curve's evolution, improves real-time forecasts of multi-quarter excess returns by a quarter to a third, and, in an affine model that lets it enter the physical dynamics but not the price, accounts for about four-fifths of the variance of the twelve-month term premium. The premium is a scarcity premium: it widens as reserves fall toward the boundary of the ample region and narrows as they are added. The overnight market shows why. A given reserve drain moves the funds rate only when reserves are scarce, so scarcity is a state in which rate risk is larger, and because the state is persistent, the futures curve prices the state while ignoring the drains themselves.

The result places reserve adequacy among the unspanned factors of the term-structure literature, with an institutional reason for its unspannedness rather than an empirical regularity, and it gives the premium that moves with reserve scarcity a market price in the units of the policy debate: basis points of twelve-month term premium per log point of reserves relative to GDP. That price is what responds as the balance sheet moves toward or away from the steep region of reserve demand. A second completed tightening cycle would sharpen the estimate, and a mean survey path would separate the premium from expectations more cleanly than one cycle allows; what the present evidence establishes is that the compensation embedded in the market for the policy rate itself moves with reserve adequacy, and that it is larger the closer the system sits to scarcity.

References

Adrian, T., R. K. Crump, and E. Moench (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics 110*(1), 110–138.

Afonso, G., M. Cipriani, A. Copeland, A. Kovner, G. La Spada, and A. Martin (2020). The market events of Mid-September 2019. Staff Report 918, Federal Reserve Bank of New York.

Afonso, G., D. Giannone, G. La Spada, and J. C. Williams (2025). Scarce, abundant, or ample? A time-varying model of the reserve demand curve. *Journal of Finance*. Forthcoming.

Anbil, S., A. Anderson, and Z. Şenyüz (2021). Are repo markets fragile? Evidence from September 2019. Finance and Economics Discussion Series 2021-028, Board of Governors of the Federal Reserve System.

Ang, A. and M. Piazzesi (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics 50*(4), 745–787.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bauer, M. D. and G. D. Rudebusch (2017). Resolving the spanning puzzle in macro-finance term structure models. *Review of Finance 21*(2), 511–553.

Bauer, M. D., G. D. Rudebusch, and J. C. Wu (2012). Correcting estimation bias in dynamic term structure models. *Journal of Business & Economic Statistics 30*(3), 454–467.

Campbell, J. Y. and R. J. Shiller (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies 58*(3), 495–514.

Campbell, J. Y. and S. B. Thompson (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies 21*(4), 1509–1531.

Cieślak, A. (2018). Short-rate expectations and unexpected returns in Treasury bonds. *Review of Financial Studies 31*(9), 3265–3306.

Cieslak, A. and P. Povala (2015). Expected returns in Treasury bonds. *Review of Financial Studies 28*(10), 2859–2901.

Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics 138*(1), 291–311.

Clouse, J., S. Infante, and Z. Şenyüz (2025). Market-based indicators on the road to ample reserves. Feds notes, Board of Governors of the Federal Reserve System. January 31, 2025.

Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*(1), 138–160.

Copeland, A., D. Duffie, and Y. D. Yang (2025). Reserves were not so ample after all. *Quarterly Journal of Economics*. forthcoming.

Dai, Q. and K. J. Singleton (2000). Specification analysis of affine term structure models. *Journal of Finance 55*(5), 1943–1978.

d'Avernas, A. and Q. Vandeweyer (2024). Treasury bill shortages and the pricing of short-term assets. *Journal of Finance 79*(6), 4083–4141.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance 57*(1), 405–443.

Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance 6*(4), 379–406.

Fleckenstein, M. and F. A. Longstaff (2024). Treasury richness. *Journal of Finance 79*(4), 2797–2842.

Greenwood, R., S. G. Hanson, and J. C. Stein (2016). The Fed's dollars and the market's sense of what the Fed is thinking. *Journal of Finance 71*(4), 1753–1788.

Greenwood, R., S. G. Hanson, J. C. Stein, and A. Sunderam (2023). A quantity-driven theory of term premia and exchange rates. *Quarterly Journal of Economics 138*(3), 1015–1059.

Haddad, V. and D. Sraer (2020). The banking view of bond risk premia. *Journal of Finance 75*(5), 2465–2502.

Hamilton, J. D. (1997). Measuring the liquidity effect. *American Economic Review 87*(1), 80–97.

Hamilton, J. D. and T. Okimoto (2011). Sources of variation in holding returns for Fed funds futures contracts. *Journal of Futures Markets 31*(3), 205–229.

Hanson, S. G. and J. C. Stein (2015). Monetary policy and long-term real rates. *Journal of Financial Economics 115*(3), 429–448.

He, Z., S. Nagel, and Z. Song (2022). Treasury inconvenience yields during the covid-19 crisis. *Journal of Financial Economics 143*(1), 57–79.

Heath, D. (2019). Macroeconomic factors in oil futures markets. *Management Science 65*(9), 4407–4421.

Joslin, S., A. Le, and K. J. Singleton (2013). Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. *Journal of Financial Economics 109*(3), 604–622.

Joslin, S., M. Priebsch, and K. J. Singleton (2014). Risk premiums in dynamic term structure models with unspanned macro risks. *Journal of Finance 69*(3), 1197–1233.

Joslin, S., K. J. Singleton, and H. Zhu (2011). A new perspective on Gaussian dynamic term structure models. *Review of Financial Studies 24*(3), 926–970.

Kim, D. H. and J. H. Wright (2005). An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates. *Finance and Economics Discussion Series 2005*(33). Federal Reserve Board.

Krishnamurthy, A. and A. Vissing-Jorgensen (2011). The effects of quantitative easing on interest rates: Channels and implications for policy. *Brookings Papers on Economic Activity 2011*(2), 215–265.

Logan, L. K. (2023). Normalizing the FOMC's balance sheet. Speech, Federal Reserve Bank of Dallas, October 18, 2023.

Logan, L. K. (2026). The banking system and the demand for reserves. Remarks at the Eleventh District Banking Conference, Federal Reserve Bank of Dallas, April 2.

López-Salido, D. and A. Vissing-Jørgensen (2023). Reserve demand, interest rate control, and quantitative tightening. Working paper, Board of Governors of the Federal Reserve System. version September 18, 2023.

Ludvigson, S. C. and S. Ng (2009). Macro factors in bond risk premia. *Review of Financial Studies 22*(12), 5027–5067.

Piazzesi, M. and E. T. Swanson (2008). Futures prices as risk-adjusted forecasts of monetary policy. *Journal of Monetary Economics 55*(4), 677–691.

Sack, B. (2004). Extracting the expected path of monetary policy from futures rates. *Journal of Futures Markets 24*(8), 733–754.

Schmeling, M., A. Schrimpf, and S. A. M. Steffensen (2022). Monetary policy expectation errors. *Journal of Financial Economics 146*(3), 841–858.

Waller, C. J. (2025). Demystifying the Federal Reserve's balance sheet. Speech at the Federal Reserve Bank of Dallas, July 10.

Warsh, K. (2025). The Federal Reserve's broken leadership. The Wall Street Journal, Opinion, November 16 (print edition November 17).

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies 21*(4), 1455–1508.

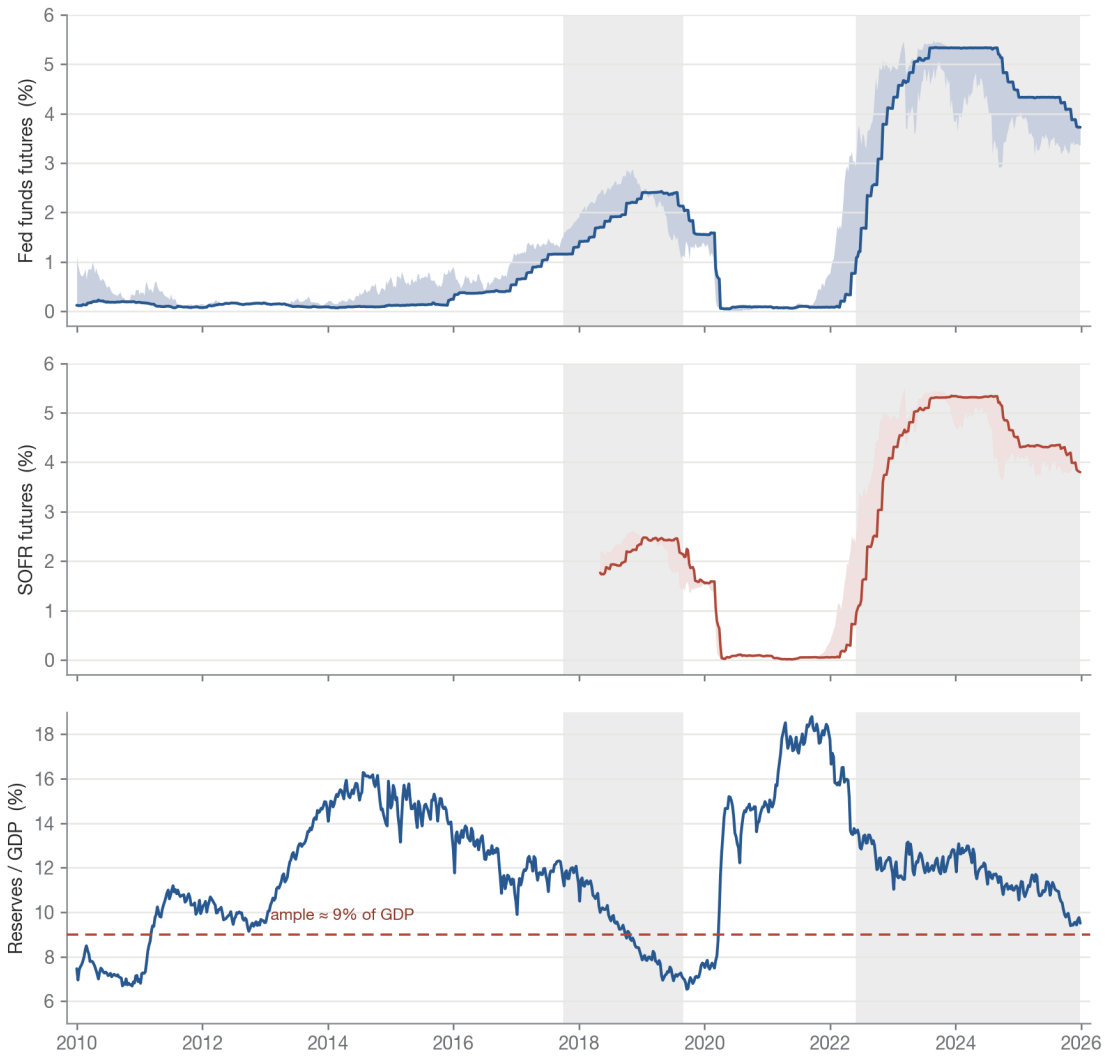


Figure 1: The two overnight-rate futures curves and reserve adequacy. The top panel plots the federal funds futures curve over the ample-reserves sample: the front contract (solid) and the range spanned by the one- to twelve-month contracts (shaded). The middle panel plots the corresponding one-month SOFR futures curve, which begins only in 2018. The bottom panel plots reserve balances as a share of nominal GDP, constructed in real time from vintage data, with the nine-percent ample threshold of Waller (2025) marked. Shaded vertical bands are the two rounds of balance-sheet runoff, 2017–2019 and 2022–2025. The September 2019 funding stress is plainly visible in SOFR and barely registers in the funds-futures curve, whose front contract moves by only a few basis points that week; reserve adequacy, by contrast, had been falling toward the ample boundary for two years.

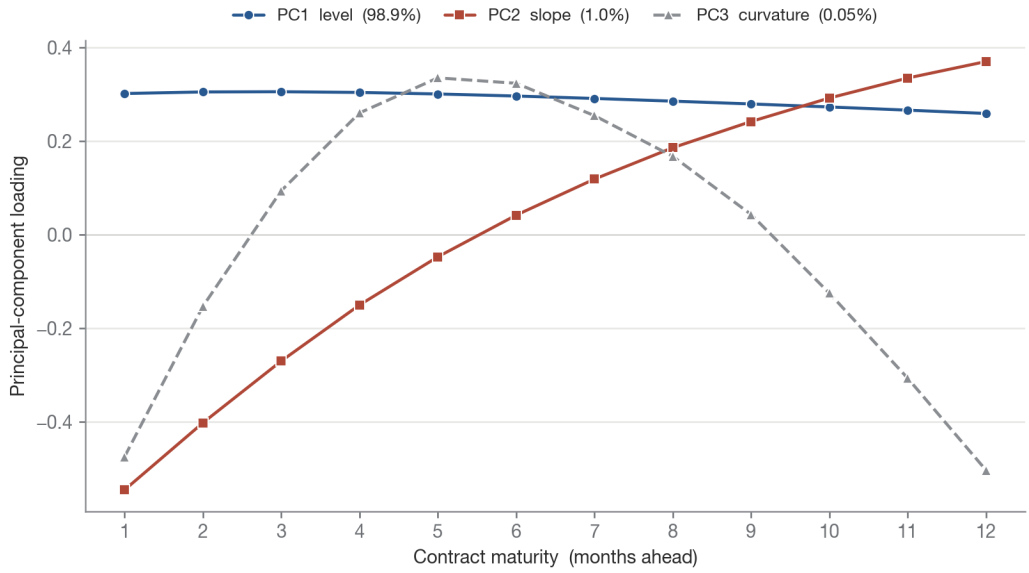


Figure 2: Principal-component loadings of the funds-futures ladder. Loadings of the first three principal components across the twelve contract maturities, ample sample. The first component loads almost uniformly (a level factor, 98.9 percent of cross-sectional variation), the second is monotone in maturity (slope, 1.0 percent), and the third is hump-shaped (curvature, 0.05 percent). The near-degeneracy leaves reserve adequacy as a candidate dimension the curve does not accommodate; three components are nonetheless retained for pricing, the minimum that fits the ends of the ladder (Section 3, Appendix D).

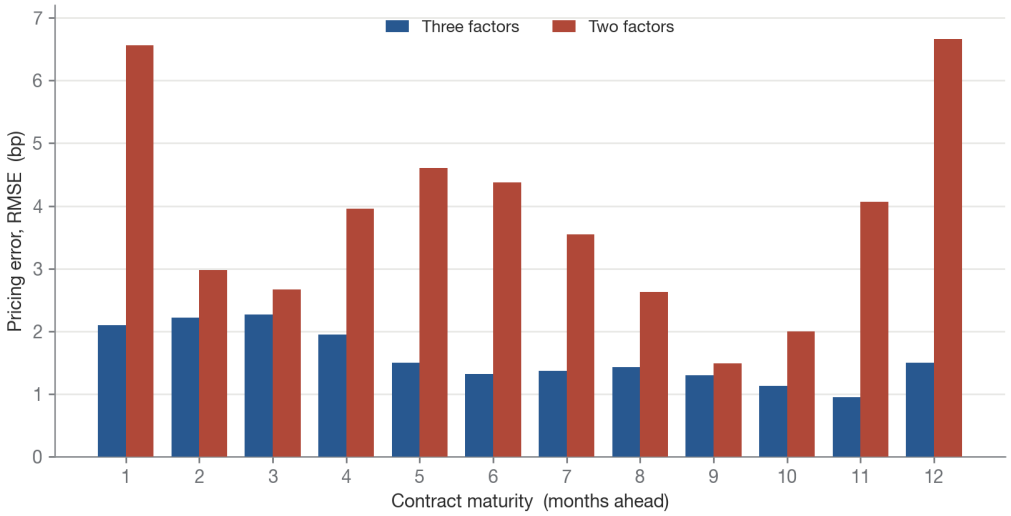


Figure 3: Cross-sectional pricing errors: two versus three factors. Root-mean-square pricing error in basis points by contract maturity, ample sample, for the three-factor pricing model (navy) and the two-factor model of Appendix D (brick). The two-factor model is comparable through the interior of the ladder but fails at both ends—6.56 basis points at one month and 6.66 at twelve—while the three-factor model stays within about two basis points throughout.

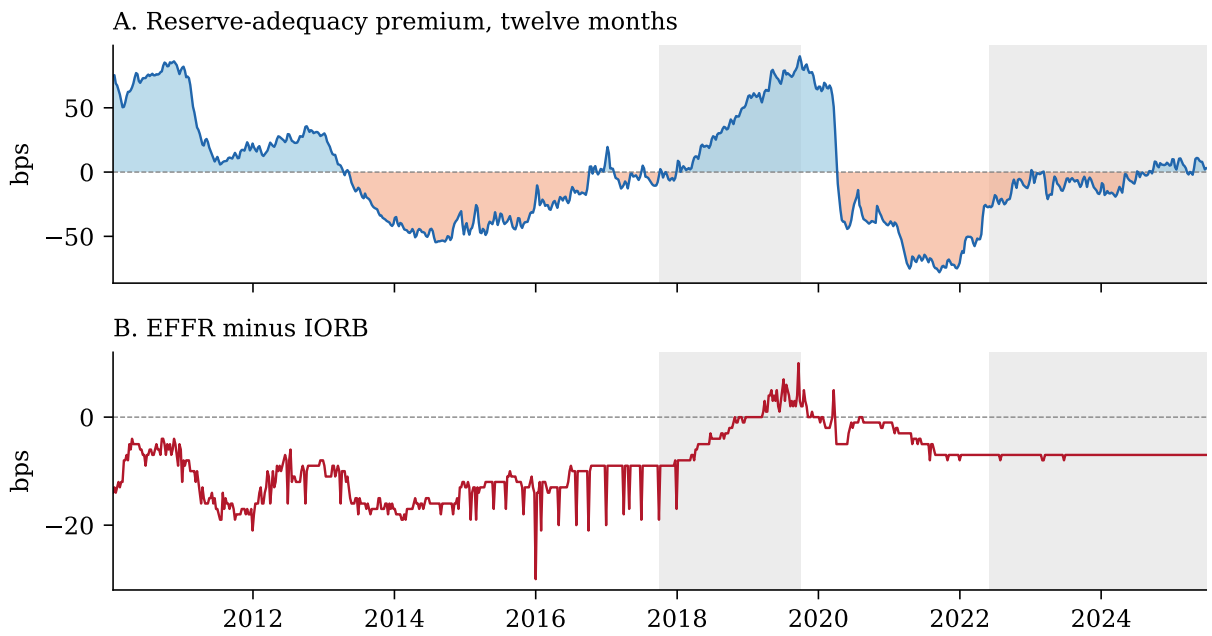


Figure 4: The premium and the scarcity dashboard. Panel A plots the reserve-adequacy premium component of the twelve-month term premium, ERP^{RA}, in basis points, computed under the least-squares physical dynamics; positive values are shaded blue and negative values red. Panel B plots the spread of the effective funds rate over the administered floor (IOER before July 2021 and IORB after), the scarcity gauge the Desk watches, in basis points. The two co-move contemporaneously, at a correlation of about 0.37. Because the gauge is a realized outcome and the premium a rescaling of the reserve series, the figure shows co-movement, not a lead; the paper makes no claim that the premium anticipates the dashboard. Shaded bands mark balance-sheet runoff, October 2017 to September 2019 and June 2022 to the end of the sample.

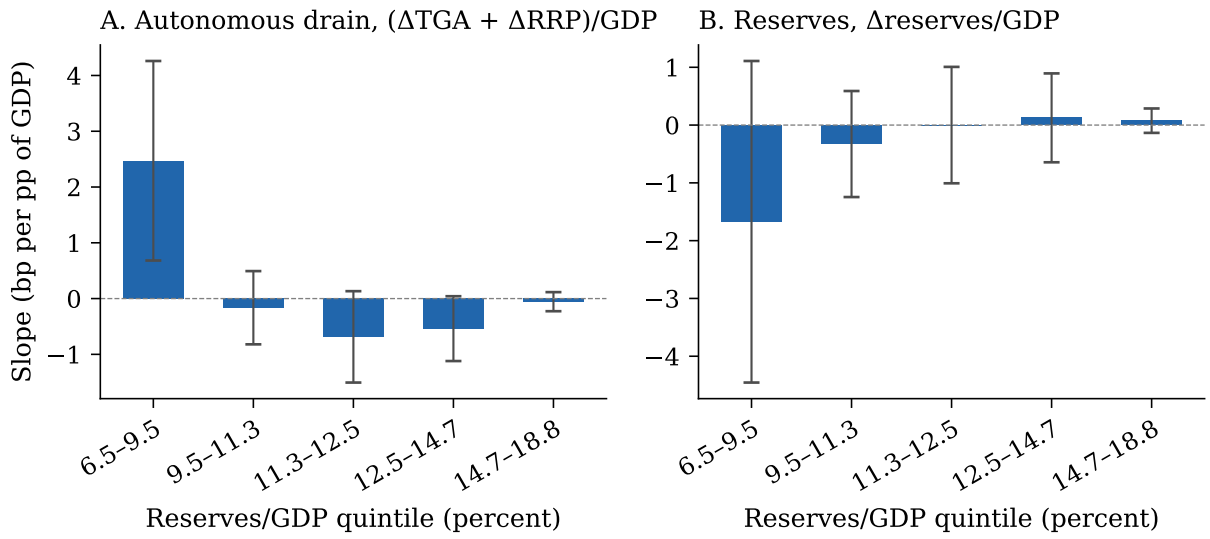


Figure 5: The overnight rate responds to reserve drains only when reserves are scarce. Panel A plots the slope of the weekly change in the spread of the effective federal funds rate over interest on reserves, Δ(EFFR − IOR) in basis points, on the weekly autonomous drain (ΔTGA + ΔRRP)/GDP in percentage points of GDP, separately for each quintile of reserves/GDP; a positive drain absorbs reserves. Panel B plots the same slopes for the weekly change in reserves/GDP, where a positive value adds reserves. The labels give the range of reserves/GDP, in percent, within each quintile. The slopes come from one pooled regression per panel with quintile-specific slopes and intercepts and common controls for target-change weeks, month-ends, quarter-ends, and the lagged spread change. Whiskers are 95 percent confidence intervals from Newey–West standard errors with four lags. Weekly (Friday) data, 8 January 2010 to 27 June 2025, 808 weeks.

Table 1: Descriptive Statistics

This table reports summary statistics for the federal funds futures panel and the reserve-adequacy factor over the ample-reserves sample, weekly (Friday) observations from January 2010 through December 2025 (835 weeks). Panel A reports the annualized implied rate ($100 - P$, in percent) on CME 30-Day Federal Funds futures by months to delivery; the current-delivery contract is excluded. The first three principal components of the ladder account for 98.91, 1.04, and 0.05 percent of its cross-sectional variation. Panel B reports reserve balances as a share of nominal GDP, built in real time from vintage data, and its demeaned logarithm m_t, the state variable used throughout (1 SD = 0.25 log points); the minimum falls in the week of 20 September 2019 and the maximum in September 2021, and the ratio ends the sample at 9.50 percent. Panel C reports futures excess returns to a long position, $rx^{(h)} = f -$ realized, in basis points; the counts are those of this 835-week descriptive panel and therefore differ from those of the balanced 808-week estimation panel used in the analysis that follows, which ends in June 2025 where the twelfth maturity ceases to be quoted. The five-quarter return uses contracts beyond the twelfth maturity under a per-observation liquidity screen (Appendix A).

	N	Mean	Std. Dev.	Min	Max
Panel A. Federal funds futures ladder (implied rate, %)					
1-month	835	1.44	1.81	0.05	5.41
2-month	835	1.46	1.81	0.05	5.45
3-month	835	1.48	1.80	0.04	5.48
4-month	835	1.48	1.78	0.04	5.48
5-month	835	1.49	1.76	0.03	5.48
6-month	835	1.49	1.73	0.03	5.45
7-month	831	1.48	1.70	0.02	5.45
8-month	827	1.47	1.67	0.00	5.40
9-month	822	1.46	1.63	0.00	5.36
10-month	818	1.46	1.60	0.00	5.31
11-month	813	1.45	1.56	−0.01	5.27
12-month	809	1.45	1.52	−0.02	5.15
Panel B. Reserve adequacy					
Reserves / GDP (%)	835	11.87	2.87	6.53	18.80
Demeaned log ratio m_t	835	0.00	0.25	−0.57	0.49
Panel C. Futures excess returns $rx^{(h)} = f -$ realized (bps)					
1 quarter	835	−0.0	20.9	−98.0	151.5
2 quarters	835	−4.4	41.0	−183.5	141.0
3 quarters	824	−8.9	64.5	−285.0	157.5
4 quarters	811	−12.8	89.9	−363.5	215.0
5 quarters	798	−15.6	115.2	−407.0	237.0

Table 2: Spanning: Reserve Adequacy Projected on the Futures Curve

This table reports R^2 from projecting reserve adequacy on the leading principal components of the twelve-contract federal funds futures ladder (equation 2) over the ample-reserves sample. "Levels" uses the series; "Innovations" uses first differences of both sides—the relevant object once the common near-unit-root trend is removed. The residual AR(1) is the weekly first-order autocorrelation of the orthogonal component $\mathrm{RA}^{\perp}$.

Components	Levels	Innovations
$K = 1$	0.007	0.002
$K = 2$	0.032	0.011
$K = 3$	0.070	0.013
Residual AR(1), $K = 3$	0.994	0.446
Observations	808	807

Table 3: Reserve Adequacy Forecasts the Curve One Period Ahead

This table reports one-week-ahead forecasting regressions of changes in the leading principal components of the federal funds futures curve on the contemporaneous curve, reserve adequacy, and three macroeconomic controls. The specification follows Heath (2019): the dependent variable is $\Delta \mathrm{PC}_{k,t+1}$ in principal-component units, and the spanned controls are the curve factors listed at the foot of each column—two principal components, three, or all twelve contract rates. The macroeconomic controls are the variables this literature uses to forecast funds-futures excess returns (Piazzesi and Swanson, 2008; Cieślak, 2018): three-month growth in nonfarm payrolls, entered with a realistic release lag; the BAA–AAA credit spread; and the VIX. Reserve adequacy and the three controls are standardized. The forecast horizon is one period, so the windows do not overlap. Reserve adequacy is the demeaned log of the two-week average of reserves over GDP, standardized; the sample is the 807 weekly transitions of the balanced panel. Newey–West standard errors with four lags are in parentheses. The third principal component, 0.05 percent of the curve's variation, is not forecastable and is omitted. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Level factor, $\Delta PC_{1,t+1}$						
Reserve adequacy	0.0310***	0.0322***	0.0289***	0.0287**	0.0282***	0.0273**
	(0.0092)	(0.0112)	(0.0100)	(0.0128)	(0.0100)	(0.0124)
Payroll growth		0.0028		0.0023		0.0011
		(0.0028)		(0.0025)		(0.0025)
Credit spread		0.0041		−0.0005		−0.0006
		(0.0123)		(0.0127)		(0.0141)
Volatility		0.0008		0.0002		−0.0051
		(0.0148)		(0.0147)		(0.0152)
Adjusted R^2 (%)	5.1	4.8	5.1	4.7	6.4	6.1
curve controls only (%)	4.1	4.1	4.3	4.3	5.6	5.6
Panel B. Slope factor, $\Delta PC_{2,t+1}$						
Reserve adequacy	0.0097**	0.0077*	0.0113***	0.0102**	0.0116***	0.0095**
	(0.0037)	(0.0041)	(0.0041)	(0.0047)	(0.0043)	(0.0047)
Payroll growth		0.0032*		0.0036**		0.0028
		(0.0017)		(0.0016)		(0.0018)
Credit spread		−0.0073		−0.0040		−0.0070
		(0.0051)		(0.0048)		(0.0048)
Volatility		0.0050		0.0055		0.0053
		(0.0050)		(0.0050)		(0.0050)
Adjusted R^2 (%)	2.1	2.0	2.3	2.1	4.2	4.1
curve controls only (%)	1.8	1.8	1.8	1.8	3.7	3.7
Curve controls	$\mathrm{PC}^{1,2}$	$\mathrm{PC}^{1,2}$	$\mathrm{PC}^{1,2,3}$	$\mathrm{PC}^{1,2,3}$	f^{1-12}	f^{1-12}
Observations	807	807	807	807	807	807

Table 4: The Reserve-Adequacy Loading in Funds-Futures Excess Returns

This table reports the predictive regression of the four-quarter federal funds futures excess return on the two leading pricing-state coordinates and reserve adequacy, with and without the standard predictors of funds-futures excess returns: three-month payroll growth, the BAA–AAA credit spread, and the VIX. The dependent variable is in percentage points. Reserve adequacy and the three return predictors are standardized; the two pricing-state coordinates enter in their own units. Newey–West standard errors, with a lag length tailored to the overlap of the forecast windows, are in parentheses. Macroeconomic series enter with realistic release lags. The table states the loading in the units of the return literature; the non-overlapping counterpart of the same conditioning exercise is Table 3. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	(1)	(2)
Reserve adequacy	-0.5295^{***}	-0.4957^{***}
	(0.1984)	(0.1695)
Payroll growth		-0.0773
		(0.0583)
Credit spread		0.1348
		(0.1027)
VIX		-0.2092^{*}
		(0.1099)
Pricing state X_1	-0.0210	-0.0002
	(0.1063)	(0.1139)
Pricing state X_2	-0.0021^{**}	-0.0021^{**}
	(0.0010)	(0.0009)
Constant	0.0050	0.0120
	(0.1935)	(0.1733)
Observations	783	783
R^2	0.418	0.471

Table 5: The Horizon Profile of Excess Returns, In and Out of Sample

The excess return is $rx^{(h)} = f -$ realized in basis points over h quarters; a positive value is compensation to a long position. Panel A regresses $rx^{(h)}$ on the first two principal components of the twelve-contract ladder and on the standardized spanning residual of reserve adequacy; the loading is in basis points per standard deviation, the incremental R^2 is the gain over the two components alone, and standard errors (in parentheses) are Newey–West with $13h$ lags. Weekly (Friday) observations, 8 January 2010 to 27 June 2025; the five-quarter return uses contracts beyond the twelfth under a liquidity screen, and the appendix tables, which apply the liquidity screen observation by observation, have 799 observations at four quarters. Panel B reports pseudo-out-of-sample forecasts made each week from January 2015 through June 2025 from least-squares regressions re-estimated on an expanding window that starts in January 2010, using only returns realized by the forecast date. The curve is the first two principal components; m is the demeaned log ratio of reserves to GDP (standard deviation 0.25); the standard predictors are twelve-month payroll growth (three-month growth in Tables 3 and 4) and the BAA–AAA credit spread, lagged one month, and the VIX. Each block compares the larger model with the nested benchmark in its heading; block (d) drops the forecasts made in 2023 and 2024. The out-of-sample R^2 is that of Campbell and Thompson (2008) against the benchmark model, in percent; CW is the one-sided Clark and West (2007) statistic with Newey–West standard errors with $13h$ lags. N is the number of observations in Panel A and the number of forecasts in blocks (a) to (c) in Panel B. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Panel A. In-sample predictive regressions

	1 quarter	2 quarters	3 quarters	4 quarters	5 quarters
Loading on reserve adequacy (bp per standard deviation)	−5.70**	−17.60***	−33.96***	−52.83***	−72.09***
	(2.68)	(5.49)	(10.17)	(16.98)	(24.28)
Incremental R^2 over the curve	0.07	0.17	0.26	0.33	0.38
N	808	808	808	808	797

Panel B. Out-of-sample forecasts

	(a) Curve + m vs. curve		(b) Curve + predictors vs. curve		(c) Curve + m + predictors vs. curve + predictors		(d) Curve + m vs. curve excl. 2023–24		
Horizon	OOS R^2 (percent)	CW t	OOS R^2 (percent)	CW t	OOS R^2 (percent)	CW t	OOS R^2 (percent)	CW t	N
1 quarter	9.1	2.64	−6.0	−1.61	11.9	2.59	1.9	2.33	548
2 quarters	25.2	2.33	−1.5	0.99	27.6	2.46	10.6	2.72	548
3 quarters	33.0	2.13	−2.5	1.10	28.6	1.80	15.5	3.09	548
4 quarters	35.4	1.92	−3.6	1.12	30.8	1.65	15.8	3.47	539

Table 6: Parameter Estimates

Panel A reports the business-day eigenvalues of the risk-neutral feedback matrix $\boldsymbol{G}$ and the risk-neutral intercept κ, estimated by concentrated quasi-maximum likelihood from the cross-section of futures rates on the balanced ample-reserves panel, with Hessian standard errors in parentheses; half-lives convert the business-day eigenvalue at 21 business days per month. The eigenvalues are ordered and separated by a floor of 10^{-4}, at which g_2 and g_3 sit (Appendix D). Panel B reports the weekly VAR(1) for the state $Z_t = (X_t', \mathrm{RA}_t)'$ (equation 6), estimated by ordinary least squares, with least-squares standard errors in parentheses; X is the pricing state inverted from the cross-section and RA is the demeaned log reserves-to-GDP ratio. The equations for X_2 and X_3 are omitted: the two states have a sample correlation of -0.9999, so only aggregates defined on the pair are estimated by the data, and how any effect divides between them is set by the eigenvalue separation floor. The reserve-adequacy column is the block through which reserve conditions reach the term premium (equation 15); its effect on the expected short rate, $\boldsymbol{\iota}'\boldsymbol{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0.0062$, is invariant to that parameterization, and the test of whether the column is zero is reported in Table 7. Under the bias-corrected dynamics of Appendix D the column's effect on the level factor is 0.022 rather than 0.041, with the same sign. Robustness of the risk-neutral block to the post-2016 subsample, to the stripped current-delivery series, and to a two-factor specification is reported in Appendix D.

Panel A. Risk-neutral block

	Estimate	Half-life (months)
g_1	1.00041	—
	(0.00013)	
g_2	0.99503	6.6
	(0.00019)	
g_3	0.99493	6.5
	(0.00020)	
κ	0.00262	—
	(0.00023)	
Correlation (g_1, κ)	-0.71	

Panel B. Physical dynamics, weekly VAR

Equation for	X_1	X_2	X_3	RA	Constant
X_1	0.9680	0.0124	0.0124	0.0414	-0.0249
	(0.0084)	(0.0071)	(0.0072)	(0.0275)	(0.0100)
RA	-0.0040	0.0015	0.0015	0.9939	-0.0006
	(0.0009)	(0.0008)	(0.0008)	(0.0031)	(0.0011)

Spectral radius of $\boldsymbol{\Phi}^{\mathbb{P}}$: 0.9924 $\quad \boldsymbol{\iota}'\boldsymbol{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0.0062$ $\quad$ Weeks: 808

Table 7: Covariance Attribution and the Test of the Reserve-Adequacy Restriction

This table reports the covariance attribution of the model-implied variance of the twelve-month term premium to reserve adequacy (equation 16), and the likelihood-ratio test of the restriction $\mathbf{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0$ that underlies it. Because the states are correlated the attribution is not an orthogonal variance share and need not lie in $[0, 1]$. The attribution under the least-squares physical dynamics is the headline estimate carried forward in the text (Section 3.4); the attribution under the bias-corrected dynamics of Bauer et al. (2012) is reported for comparison. The test compares the physical VAR with and without the reserve-adequacy column of $\mathbf{\Phi}^{\mathbb{P}}_{X,\cdot}$ and is referred to its asymptotic χ^2 distribution. The two-factor cross-check is the same restriction in the two-factor model of Appendix D.

	Value
Panel A. Covariance attribution	
Least-squares dynamics (headline)	0.79
Bias-corrected dynamics of Bauer et al. (2012) (comparison)	0.53
Panel B. Test of $\mathbf{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0$	
Likelihood-ratio statistic ($df = 3$)	11.03
$\chi^2(3)$ p-value	0.012
Two-factor cross-check: likelihood-ratio statistic ($df = 2$)	12.80
$\chi^2(2)$ p-value	0.002

Table 8: Horizon Profile and Over-Identification

This table reports, by horizon, the reserve-adequacy loading of the model-implied term premium and the corresponding reduced-form predictive slope, both in the same $rx = f -$ realized convention (so both are positive) and in basis points per one-log-point *decline* in reserve adequacy. One standard deviation of the state is 0.254 log points, so a one-standard-deviation *decline* in reserve adequacy raises the model-implied premium by 4.6, 14.1, 26.7 and 40.6 basis points at one to four quarters, the figures in the last row. The reduced-form slope is the loading of Panel A of Table 5 with its sign reversed and divided by the standard deviation of the state; its standard errors are reported there. The two are shown together as a consistency check on the structural estimate, not as a second significance test. The last row is the standard deviation of the reserve-adequacy premium component at each horizon, equal to the model loading times the sample standard deviation of the state.

	Horizon (h quarters)			
	1	2	3	4
Model premium loading (bps)	18.0	55.4	105.1	159.8
Reduced-form slope $\hat{\gamma}$ (bps)	22.4	69.2	133.5	207.7
Premium share (model / $\hat{\gamma}$)	0.80	0.80	0.79	0.77
ERP^{RA} std. dev. (bps)	4.6	14.1	26.7	40.6

Table 9: The Reserve-Adequacy Premium across Balance-Sheet Regimes

This table reports the mean of the model-implied reserve-adequacy premium component of the twelve-month term premium, ERP^{RA}, over four non-overlapping balance-sheet regimes; N is the number of weeks. The regimes are dated windows and do not exhaust the sample: the weeks of transition between them are omitted. By construction the series moves one-for-one with reserve adequacy (its correlation with m_t is -1), so the table describes how reserve conditions traversed the sample, not an independent signal. Over the full sample ERP^{RA} reaches +90.1 basis points in the weeks around the September 2019 repo spike and -78.0 basis points in September 2021; it co-moves with the spread of the funds rate over the administered floor at a correlation of 0.37. In the common $rx = f -$ realized convention the correlation between reserve adequacy and the full twelve-month term premium is -0.91 (scarce reserves, high premium).

Regime	N (weeks)	Mean ERP^{RA} (bps)
Early ample (2010–2017)	404	+1.2
First tightening (2017–2019)	100	+36.4
Pandemic QE (2020–2021)	96	−47.1
Second tightening (2022–2025)	161	−6.4

Table 10: Reserve Flows, Reserve Levels, and the Overnight Rate

Panel A: weekly (Friday) data, 8 January 2010 to 27 June 2025, 808 weeks. The dependent variable is the weekly change in the spread of the effective federal funds rate over interest on reserves, Δ(EFFR − IOR), in basis points. Regressors are weekly changes in H.4.1 items in percent of GDP; a positive drain (TGA or reverse repurchase agreements) absorbs reserves. LowRA equals one when reserves/GDP, denoted RA, is below 10 percent. Each regression interacts the regressor with LowRA and includes LowRA, indicators for target-change weeks, month-ends, and quarter-ends, and the lagged spread change. Column (1) adds changes in total Federal Reserve assets/GDP and currency/GDP, each interacted with LowRA; column (2) enters ΔTGA/GDP jointly with ΔRRP/GDP and its interaction, not reported; column (3) uses the sum of the two drains. Newey–West standard errors, four lags. Panel B: daily data. The spot row is the one-day change in EFFR−IOR; the futures rows pool one-day changes in the implied rates of contracts in the stated range of months to delivery, with open interest of at least 1,000 contracts and contract fixed effects. Events are tax dates, mid-month Treasury settlements, and month-ends. The second row of each block is the additional effect when RA is below 10 percent. Controls are a target-change indicator and LowRA. Standard errors are Newey–West with five lags (spot) or clustered by date (futures). Panel C: one-week-ahead change in the first or second principal component of the twelve-contract ladder, times 100, on the two current components and reserve adequacy m, the demeaned log ratio of reserves to GDP (one standard deviation is 0.25), alone or with its one-week or four-week change; state variables are standardized; 807 weeks. Newey–West standard errors, four lags. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Panel A. Weekly spread changes on reserve flows, dependent variable Δ(EFFR − IOR)

	(1) ΔReserves/GDP	(2) ΔTGA/GDP	(3) (ΔTGA+ΔRRP)/GDP
Slope, reserves/GDP ≥ 10 percent	0.31*	−0.31	−0.35**
	(0.17)	(0.30)	(0.16)
Slope, reserves/GDP < 10 percent	−2.62***	1.83***	2.22***
	(0.78)	(0.66)	(0.72)
Difference (× LowRA)	−2.94***	2.14***	2.57***
	(0.79)	(0.68)	(0.72)
R^2	0.265	0.171	0.167

Panel B. Daily one-day changes on reserve-drain event days (bp)

	Tax dates	Mid-month	Month-end
Spot EFFR−IOR, RA ≥ 10 percent	0.06	0.00	−2.72***
3,887 days	(0.08)	(0.05)	(0.34)
Additional, RA < 10 percent	2.35**	1.01***	2.19***
	(1.04)	(0.32)	(0.54)
Futures 1–3, RA ≥ 10 percent	−0.13	−0.10	−0.15*
11,611 contract-days	(0.16)	(0.14)	(0.08)
Additional, RA < 10 percent	0.10	0.15	−0.20
	(0.26)	(0.22)	(0.48)
Futures 4–6, RA ≥ 10 percent	−0.32	−0.16	−0.30
11,612 contract-days	(0.56)	(0.30)	(0.20)
Additional, RA < 10 percent	0.21	0.13	−0.45
	(0.72)	(0.42)	(0.62)
Futures 7–12, RA ≥ 10 percent	−0.33	−0.22	−0.65*
23,208 contract-days	(0.70)	(0.45)	(0.33)
Additional, RA < 10 percent	0.09	−0.15	−0.56
	(0.97)	(0.60)	(0.72)

Panel C. One-week-ahead curve changes on the level and flow of reserve adequacy

	$\Delta PC_{1,t+1} \times 100$			$\Delta PC_{2,t+1} \times 100$		
	(1)	(2)	(3)	(1)	(2)	(3)
Reserve adequacy m_t (level)	3.12***	3.11***	3.11***	1.01***	0.96**	1.04***
	(0.93)	(0.92)	(0.93)	(0.38)	(0.38)	(0.38)
Δm_t (one-week flow)		0.21			0.99*	
		(1.02)			(0.56)	
$\Delta_4 m_t$ (four-week flow)			0.05			−0.20
			(0.85)			(0.42)
R^2	0.055	0.055	0.055	0.025	0.030	0.026

A Data Construction

This appendix describes the construction of the futures panel and the reserve-adequacy factor used throughout the paper.

A.1 Contract-Level Data and Validation

The futures data are contract-level daily settlement prices, volumes, and open interest for CME 30-Day Federal Funds futures, covering 354 consecutive delivery months from January 1997 to June 2026 (204,000 contract-days); every delivery month in the sample period is present. Because the vendor record reuses contract identifiers across cycles, each contract is validated against four gates before it enters the panel. The *level* gate compares the final settlement price to 100 minus the arithmetic average of the effective funds rate over the labeled delivery month; the *discrimination* gate requires the labeled month to fit better than its neighbors by a margin exceeding the resolution of the data; the *argmin* gate requires the best-fitting month to be the labeled one; and the *coverage* gate requires the contract to be quoted over the window in which it is used. On the estimation sample the level gate passes with a median absolute error of 0.30 basis points in 2002–2008, 0.033 in 2009–2015, and 0.016 in 2016–2025. During the zero-lower-bound years the discrimination gate is degenerate—with the funds rate pinned near zero, neighboring delivery months are genuinely indistinguishable—and those contracts are marked as undiscriminated rather than passed. The 1997–2001 contracts lack final settlement rows and are never used.

A.2 The Weekly Panel and the Pricing Ladder

The panel is sampled on Fridays, with the preceding trading day substituted on holidays, and contracts are indexed by months to delivery. The contract in its delivery month is excluded, since its rate is mechanically pinned by the funds rate already accrued within the month; it re-enters the affine model only in the stripped form of Section 3.4. A liquidity

screen counts a contract as tradable in a week if its open interest is at least one thousand contracts. The pricing ladder used by the affine model is then truncated at the deepest maturity that clears the screen in at least 95 percent of the ample-regime weeks: the twelfth maturity is liquid in 96 percent of weeks and the thirteenth in 94 percent, so the ladder is twelve deep.

A.3 Excess Returns and the Two Panels

The h-quarter excess return is computed with both legs on the contract's own accrual window: the rate implied by the quoted price at date t, and the settlement rate the same contract eventually pays (equation 3). Because the ladder is monthly and the horizons are quarterly, the realized holding period differs from its nominal target by less than half a day on average at every horizon; observations whose accrual window would deviate by more than 45 days are dropped.

The return panel and the pricing ladder are screened differently, for different reasons, and it is worth being explicit about why. The affine model requires a *balanced* cross-section, because the pricing state is inverted from a fixed set of contracts present in every week; this is the ladder truncated at twelve. A predictive regression, by contrast, requires only a tradable quote at the date the return begins, so the return panel is unbalanced and applies the liquidity screen observation by observation. The five-quarter return accordingly draws on contracts beyond the twelfth: its modal contract is the fifteenth, its range the sixth through the sixteenth, the contract used is liquid in 88 percent of weeks, 4.4 percent of observations sit at the 45-day accrual cap, and 798 dates survive. No quantity of the affine model uses these longer contracts.

A.4 Reserve Adequacy in Real Time

Reserve adequacy is reserve balances scaled by nominal GDP, constructed so that its value at date t uses only information available at t. The numerator is total reserve balances from the Federal Reserve's H.4.1 release, a Wednesday stock that enters the weekly grid with a two-business-day publication delay and is averaged over two weeks to match the reserve maintenance period. The denominator is nominal GDP taken from the ALFRED real-time archive at its *first release* and dated by its actual publication date, rather than the later revised value. Within the ample regime the logarithm of the ratio is demeaned; that demeaned log ratio is the state variable RA_t, whose standard deviation is 0.25 log points. The scarce-reserve regime of the 2000s is excluded from the analysis for the reasons given in Section 2.1.

B Mathematical Derivations

The pricing kernel and the change of measure

The model is closed by an essentially affine change of measure (Duffee, 2002). The one-period density of $\mathbb{Q}$ with respect to $\mathbb{P}$ is

$$\xi_{t,t+1} \;=\; \exp\Big(-\tfrac{1}{2}\,\lambda_t'\lambda_t - \lambda_t'\,\varepsilon_{t+1}\Big), \qquad \lambda_t = \boldsymbol{\Lambda}_0 + \boldsymbol{\Lambda}_1 Z_t, \tag{B.1}$$

where ε_{t+1} are the standardized innovations of the physical VAR (6) and λ_t is the market price of risk. A futures price is a martingale under $\mathbb{Q}$, so no discounting enters its pricing and the change of measure acts only through the conditional mean of the state,

$$\boldsymbol{\Phi}^{\mathbb{Q}} \;=\; \boldsymbol{\Phi}^{\mathbb{P}} - \boldsymbol{\Sigma}_Z\,\boldsymbol{\Lambda}_1, \qquad \boldsymbol{\mu}^{\mathbb{Q}} \;=\; \boldsymbol{\mu}^{\mathbb{P}} - \boldsymbol{\Sigma}_Z\,\boldsymbol{\Lambda}_0, \tag{B.2}$$

with $\boldsymbol{\Sigma}_Z$ the Cholesky factor of the innovation covariance. Restricting the pricing factors so that reserve adequacy has no column in $\boldsymbol{\Phi}^{\mathbb{Q}}$, $[\boldsymbol{\Phi}^{\mathbb{Q}}]_{X,\cdot} = [\boldsymbol{G} \;\; 0]$, and does not enter the short rate gives the risk-neutral block (8): futures prices depend on X_t alone, and because $\boldsymbol{\Sigma}_Z$ is lower triangular with the pricing factors ordered first, the pricing-factor rows of $\boldsymbol{\Lambda}_1$ are $\boldsymbol{\Sigma}_X^{-1}(\boldsymbol{\Phi}^{\mathbb{P}}_{X,\cdot} - [\boldsymbol{G} \;\; 0])$ as in (9), and their reserve-adequacy column is exactly $\boldsymbol{\Sigma}_X^{-1}\boldsymbol{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}}$—entirely physical, as stated in Section 3.1.

In estimation the two blocks are fitted at different frequencies (Section 3.4): the risk-neutral block at the business-day step of the model and the physical VAR on weekly observations, so that the estimated $\boldsymbol{\Phi}^{\mathbb{P}}$ is a one-week transition. A comparison of the two at a common step would set the one-week risk-neutral transition of the pricing factors to $\boldsymbol{G}^5$, five business-day steps of (8). No reported quantity requires that comparison: the estimation never forms λ_t, and the term premium is the difference of the two expectations in (13), each averaged over the same days of the delivery month, with the physical leg iterated as described below.

Proof of Proposition 1

Under (8) the risk-neutral conditional mean of the state h business days after the quotation date t is

$$\mathbb{E}_t^{\mathbb{Q}}[X_{t+h}] \;=\; \boldsymbol{G}^h X_t + \sum_{j=0}^{h-1} \boldsymbol{G}^j \, \boldsymbol{\mu}^{\mathbb{Q}},$$

and since $\boldsymbol{G}$ is diagonal and $\boldsymbol{\mu}^{\mathbb{Q}} = (\kappa, 0, 0)'$ is zero outside its first coordinate, $\sum_{j=0}^{h-1} \boldsymbol{G}^j \boldsymbol{\mu}^{\mathbb{Q}} = (\kappa S_h, 0, 0)'$; applying $\boldsymbol{\iota}'$ to both sides gives the risk-neutral expectation of the short rate,

$$\mathbb{E}_t^{\mathbb{Q}}[r_{t+h}] \;=\; \boldsymbol{\iota}'\boldsymbol{G}^h X_t + \kappa S_h, \qquad S_h = \sum_{j=0}^{h-1} g_1^j. \tag{B.3}$$

The contract for delivery month m settles on $\sum_{h \in \mathcal{H}_m} \omega_{m,h}\, r_{t+h}$, an affine function of the states on the business days that set the month's rates. A 30-day federal funds futures

pays 100 minus that average, and under $\mathbb{Q}$ the futures price is a martingale, so the quoted rate equals the risk-neutral expectation of the settlement average; because the settlement rate is *affine* in the underlying (an arithmetic average, not a compounded one), there is no convexity term, and the wedge induced by daily marking-to-market is of second order at these horizons (Piazzesi and Swanson, 2008). Taking $\mathbb{E}_t^{\mathbb{Q}}$ through the average,

$$f_t^{(m)} \;=\; \sum_{h\in\mathcal{H}_m} \omega_{m,h}\,\mathbb{E}_t^{\mathbb{Q}}\big[r_{t+h}\big] \;=\; \sum_{h\in\mathcal{H}_m} \omega_{m,h}\Big(\boldsymbol{\iota}'\boldsymbol{G}^{\,h}X_t + \kappa\, S_h\Big),$$

by (B.3); collecting the state-dependent and constant terms gives $\bar{B}_t^{(m)}$ of (11) and $\bar{A}_t^{(m)}$ of (12). ∎

The term-premium decomposition

Iterating the physical VAR (6) over w steps gives $\mathbb{E}_t^{\mathbb{P}}[Z_{t+w}] = \Big(\sum_{j=0}^{w-1}(\boldsymbol{\Phi}^{\mathbb{P}})^j\Big)\,\boldsymbol{\mu}^{\mathbb{P}} + (\boldsymbol{\Phi}^{\mathbb{P}})^w Z_t$. In the model one step is one business day and $w = h$; with $\boldsymbol{\Phi}^{\mathbb{P}}$ estimated as a one-week transition, a horizon of h business days is iterated over $w(h) = \mathrm{round}(h/5)$ weekly steps, which is the form used in computation. Writing $[\,\cdot\,]_X$ and $[\,\cdot\,]_{X,\cdot}$ for the rows of a vector or matrix on the full state that belong to the three pricing factors,

$$\mathbb{E}_t^{\mathbb{P}}\left[\sum_{h\in\mathcal{H}_m} \omega_{m,h}\, r_{t+h}\right] \;=\; \sum_{h\in\mathcal{H}_m} \omega_{m,h}\left[\boldsymbol{\iota}'\Big[\Big(\sum_{j=0}^{w(h)-1}(\boldsymbol{\Phi}^{\mathbb{P}})^j\Big)\boldsymbol{\mu}^{\mathbb{P}}\Big]_X + \boldsymbol{\iota}'\big[(\boldsymbol{\Phi}^{\mathbb{P}})^{w(h)}\big]_{X,\cdot}\, Z_t\right].$$

Subtracting this from (10) yields (14), with the reserve-adequacy loading (15) obtained by selecting the RA coordinate. Because reserve adequacy enters the futures rate with a zero loading, the entire reserve-adequacy loading of the term premium comes from the physical leg; and because $(\boldsymbol{\Phi}^{\mathbb{P}})^w$ accumulates, a persistent RA produces a loading that grows with the delivery horizon.

The variance share and the specification test

The estimand of Section 4 is the fraction of the model-implied variance of the twelve-month term premium attributable to reserve adequacy. Writing $\mathrm{TP}_t^{(12)} = \beta_0 + (\beta_X)' X_t + \beta_{\mathrm{RA}} \mathrm{RA}_t$ from (14), the share of equation (16) is

$$\mathrm{share}_{\mathrm{RA}} \;=\; \frac{\beta_{\mathrm{RA}}\,\mathrm{Cov}(\mathrm{RA}_t,\, \mathrm{TP}_t^{(12)})}{\mathrm{Var}(\mathrm{TP}_t^{(12)})},$$

a covariance share (so the curve and reserve-adequacy shares need not sum to one). It is a scalar defined on the whole state and is therefore invariant to the rotation of the near-degenerate curve subspace discussed in Section 3.6. The associated hypothesis is $H_0 : \boldsymbol{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0$, the joint restriction that reserve adequacy does not enter the physical dynamics of any pricing factor; under H_0 the loading $\beta_{\mathrm{RA}}^{(m)}$ of (15) is identically zero at every horizon. Its likelihood-ratio statistic has three degrees of freedom (one per pricing factor) and is referred to the asymptotic $\chi^2(3)$ distribution; the value is 11.03 (Table 7).

C A Shadow-Rate Counterfactual

The Gaussian model of Section 3 permits negative short rates, and two episodes in the estimation window—2010–2015 and 2020–2021—sit at the zero lower bound. This appendix asks how much of the reserve-adequacy component of the term premium survives when the short rate is floored.

The exercise is a counterfactual, not a re-estimation. We hold the estimated dynamics fixed and replace the short rate with $r_t = \max(\underline{r},\, \boldsymbol{\iota}' X_t)$, taking $\underline{r} = 10$ basis points in the main version and zero in a robustness version.

Because the floor is applied inside an expectation, the pricing formula of Proposition 1 is replaced by an expectation of a floored average. For a day of the delivery month at horizon h, with weekly step $w = w(h)$, the short rate $r_{t+h} = \boldsymbol{\iota}' X_{t+h}$ is evaluated at the

weekly state, which is Gaussian under $\mathbb{P}$ with

$$\mu_{t,w} = \boldsymbol{\iota}'\Big[\Big(\textstyle\sum_{j=0}^{w-1}(\boldsymbol{\Phi}^{\mathbb{P}})^j\Big)\boldsymbol{\mu}^{\mathbb{P}} + (\boldsymbol{\Phi}^{\mathbb{P}})^w X_t\Big], \qquad \sigma_w^2 = \boldsymbol{\iota}'\, V_w\, \boldsymbol{\iota}, \quad V_w = \textstyle\sum_{j=0}^{w-1}(\boldsymbol{\Phi}^{\mathbb{P}})^j\, \boldsymbol{\Sigma}\boldsymbol{\Sigma}'\, \big((\boldsymbol{\Phi}^{\mathbb{P}})^j\big)'. \tag{C.1}$$

The truncated-Gaussian mean gives, for each day,

$$\mathbb{E}_t^{\mathbb{P}}\big[\max(\underline{r},\, r_{t+h})\big] \;=\; \underline{r}\,\Phi(d_w) + \mu_{t,w}\big(1-\Phi(d_w)\big) + \sigma_w\,\phi(d_w), \qquad d_w = \frac{\underline{r}-\mu_{t,w}}{\sigma_w}, \tag{C.2}$$

with ϕ, Φ the standard normal density and cdf. Weighting (C.2) at $w = w(h)$ by $\omega_{m,h}$ and summing over the delivery month gives the floored expected settlement average,

$$\mathbb{E}_t^{\mathbb{P}}\Big[\textstyle\sum_{h\in\mathcal{H}_m}\omega_{m,h}\max(\underline{r},\, r_{t+h})\Big] \;=\; \sum_{h\in\mathcal{H}_m}\omega_{m,h}\Big[\underline{r}\,\Phi(d_{w(h)})+\mu_{t,w(h)}\big(1-\Phi(d_{w(h)})\big)+\sigma_{w(h)}\phi(d_{w(h)})\Big]. \tag{C.3}$$

Equation (C.3) is the floored counterpart of the physical leg of the term premium (13); the risk-neutral (futures) leg does not depend on reserve adequacy, so in the *difference* that defines the reserve-adequacy component it cancels, and (C.3) is all that must be recomputed. The reserve-adequacy component under the floor is thus

$$\Delta\mathrm{TP}^{\mathrm{RA}}_{\text{floored}} \;=\; \mathbb{E}_t^{\mathbb{P}}\big[\mathrm{avg}_{\text{floored}} \mid \mathrm{RA}_{\mathrm{hi}}\big] - \mathbb{E}_t^{\mathbb{P}}\big[\mathrm{avg}_{\text{floored}} \mid \mathrm{RA}_{\mathrm{lo}}\big], \tag{C.4}$$

to be compared with the affine loading $\beta^{(12)}_{\mathrm{RA}}\,\Delta\mathrm{RA}$ of (15).

Differentiating (C.2), $\partial\,\mathbb{E}[\max(\underline{r}, r)]/\partial\mu = 1-\Phi(d_w) = \mathbb{P}(r > \underline{r})$. The floored per-day reserve-adequacy loading is therefore the affine per-day loading multiplied by $1-\Phi(d_w)$, averaged over the delivery month with the weights $\omega_{m,h}$. When the funds rate is well above the floor $\Phi(d_w) \to 0$ and the loading is unchanged; as the rate approaches the bound $\Phi(d_w)$ rises and the loading shrinks. This is exactly the pattern in Table C.1: the standard-deviation ratio is 1.00 above 2.5 percent, 0.91 in the one-to-two-and-a-half percent band, and 0.63 below 25 basis points, and the full-sample share falls from 0.79 to 0.69

(ten-basis-point floor) or 0.71 (zero floor). The construction is specific to the unspanned specification, in which RA does not enter the pricing side; it delivers differences in the reserve-adequacy component, not the level of a shadow-rate term premium, which we do not report.

Table C.1 reports the result by the level of the policy rate. When the funds rate is above 2.5 percent—the later part of the first tightening episode and most of the second—the floored and affine reserve-adequacy components are the same series to within a thousandth of a basis point, a standard-deviation ratio of 1.00. Below that the ratio is 0.91 in the one-to-two-and-a-half percent range, which covers the 2017 to mid-2018 stretch of the first tightening episode, and 0.63 in weeks with the funds rate below 25 basis points, where a floored short rate mechanically removes the rate risk the premium compensates. Over the full sample the reserve-adequacy share of the term-premium variance falls from 0.79 to 0.69 under a ten-basis-point floor and to 0.71 with a zero floor, and the decline comes from the zero-lower-bound weeks. The floored shares are computed against the affine price variance in the denominator, so they are comparable to one another and to the affine share but are not a self-contained decomposition of a shadow-rate model; the standard-deviation ratios are the cleaner object and are what we rely on.

Table C.1: Shadow-Rate Counterfactual, by the Level of the Policy Rate

Ratio of the standard deviation of the reserve-adequacy component of the twelve-month term premium under a floored short rate to the same object under the affine model, holding all estimated dynamics fixed; buckets are defined by the level of the effective funds rate in the week. The floor is 10 basis points. The share of the term-premium variance attributable to reserve adequacy is 0.79 under the affine model, 0.69 under the 10-basis-point floor, and 0.71 under a zero floor, the floored versions being measured against the affine price variance.

Weeks with the effective funds rate	Weeks	SD ratio (floored / affine)	Max. abs. diff. (bp)
Below 0.25 percent	415	0.626	58.9
0.25 to 1 percent	90	0.960	5.6
1 to 2.5 percent	158	0.913	46.1
Above 2.5 percent	145	1.000	0.000
All weeks	808	0.75	

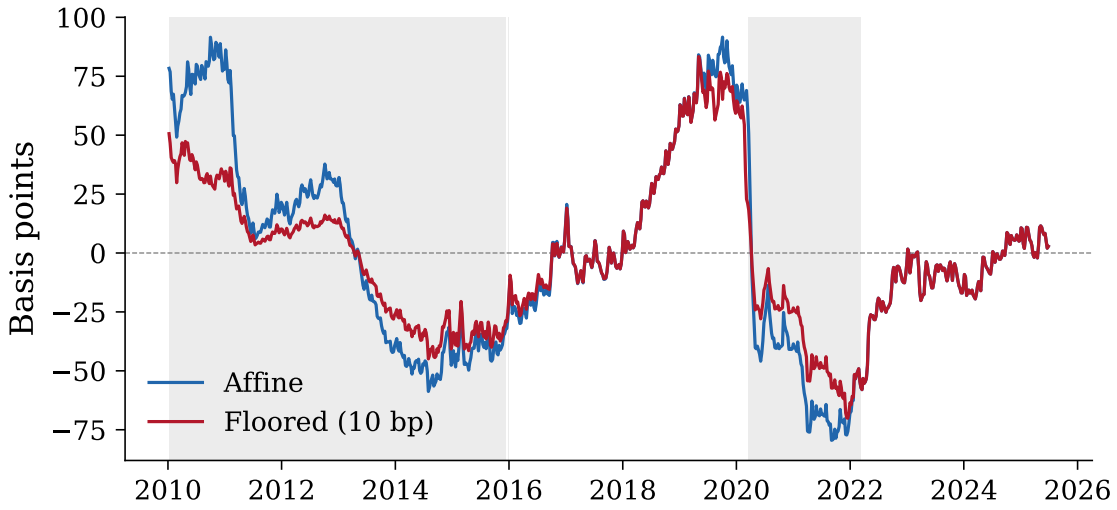


Figure C.1: The reserve-adequacy component of the term premium under the affine model and under a floored short rate. The two series are plotted over the ample sample; shaded bands mark weeks at the zero lower bound. The series nearly coincide away from the bound and compress within the bands.

D Estimation Details and Specification Checks

D.1 The Estimator

The model is estimated in two steps, the risk-neutral block first and the physical dynamics second.

Each of the twelve contract series carries an independent Gaussian pricing error,

$$f_t^{\text{obs},m} = f_t^{(m)} + \nu_{m,t}, \qquad \nu_{m,t} \sim \mathcal{N}(0, \zeta_m^2), \tag{D.1}$$

where $f_t^{(m)}$ is the affine rate of Proposition 1 and the three portfolios $P_t = \boldsymbol{W} f_t$ are priced exactly. Concentrating out the ζ_m, the risk-neutral parameters minimize the quasi-log-likelihood

$$\hat{\theta}_{\mathbb{Q}} = \underset{g_1, g_2, g_3, \kappa}{\arg\min} \sum_{m=1}^{12} T_m \log \hat{\zeta}_m^2(\theta_{\mathbb{Q}}), \qquad \hat{\zeta}_m^2 = \frac{1}{T_m} \sum_t \hat{\nu}_{m,t}^2, \tag{D.2}$$

from 100 starting values on a coarse grid with a frozen seed; the eigenvalues are ordered

and separated by a floor of 10^{-4}. Two conventions are declared rather than defended: with three portfolios priced exactly the twelve residuals lie in a nine-dimensional space, so the criterion is a quasi-likelihood; and the Jacobian of the change of variables from prices to states is omitted, as in Heath (2019). The current-delivery contract is excluded from the twelve and reintroduced as a thirteenth measurement series, its accrued portion stripped from the published fixings and the remainder priced by the same loadings restricted to the $N_{\mathrm{rem},t}$ days still to accrue, with a measurement standard deviation scaled by $N_m/N_{\mathrm{rem},t}$; the quarter-end months of 2018–2019 are excluded by a calendar rule.

Given $\hat{\theta}_{\mathbb{Q}}$, the pricing state is inverted from the cross-section date by date,

$$\hat{X}_t \;=\; (\boldsymbol{W}\bar{B}_t)^{-1}\big(P_t - \boldsymbol{W}\bar{A}_t\big)\,, \tag{D.3}$$

and the weekly first-order VAR for $\hat{Z}_t = (\hat{X}_t', \mathrm{RA}_t)'$ is estimated by ordinary least squares,

$$\hat{Z}_{t+1} \;=\; \boldsymbol{\mu}^{\mathbb{P}} + \boldsymbol{\Phi}^{\mathbb{P}}\hat{Z}_t + \boldsymbol{\Sigma}_Z\,\varepsilon_{t+1}, \tag{D.4}$$

where one step is one week of observations. The VAR is run in the inverted coordinates rather than the observed portfolios P_t because the loadings $\bar{B}_t$ move with the calendar within the delivery month; only in the state coordinates is $\boldsymbol{\Phi}^{\mathbb{P}}$ time-invariant.

Over a single balance-sheet cycle with a near-unit-root state, least squares understates the autoregressive coefficients of (D.4), and the downward bias propagates into the reserve-adequacy loading (15), where it compounds with horizon. We correct it following Bauer et al. (2012). The bias

$$\hat{\boldsymbol{B}} \;=\; \mathbb{E}\Big[\hat{\boldsymbol{\Phi}}^{\mathbb{P}}{}_{\mathrm{ols}} \;\Big|\; \boldsymbol{\Phi}^{\mathbb{P}} = \hat{\boldsymbol{\Phi}}^{\mathbb{P}}{}_{\mathrm{ols}}\Big] - \hat{\boldsymbol{\Phi}}^{\mathbb{P}}{}_{\mathrm{ols}} \tag{D.5}$$

is estimated by simulating paths from the fitted VAR and averaging the least-squares estimator across them, and the corrected transition matrix is

$$\hat{\boldsymbol{\Phi}}^{\mathbb{P}}{}_{\mathrm{bc}} \;=\; \hat{\boldsymbol{\Phi}}^{\mathbb{P}}{}_{\mathrm{ols}} - \hat{\boldsymbol{B}}. \tag{D.6}$$

When $\hat{\boldsymbol{\Phi}}^{\mathbb{P}}_{\text{bc}}$ is explosive, the stationarity adjustment of Bauer et al. (2012) shrinks the correction along $\hat{\boldsymbol{\Phi}}^{\mathbb{P}}_{\text{ols}} - \delta\hat{\boldsymbol{B}}$, reducing δ from one until the largest eigenvalue falls below unity.

Standard errors for the risk-neutral block are Hessian-based. The test of $\boldsymbol{\Phi}^{\mathbb{P}}_{X,\text{RA}} = 0$ in Section 4 is a likelihood-ratio test of the physical VAR with and without the reserve-adequacy column, referred to its asymptotic χ^2 distribution, as in Joslin et al. (2014).

D.2 Results and Specification Checks

This subsection reports the estimates and the three limitations flagged in Section 3.6.

Table 6 reports the risk-neutral block. The second and third factors decay with a half-life of about six and a half months and the first is effectively a random walk; all 100 starting values reach the same optimum. As an out-of-model consistency check, the implied overnight rate $\boldsymbol{\iota}'\hat{X}_t$ tracks the realized effective funds rate, a series the estimation never targets, with a median absolute deviation of a few basis points, and the larger deviations are confined to the March 2020 collapse and the 2022 liftoff.

The reserve-adequacy attribution is 0.79 under least squares and 0.53 after the correction. The correction raises the estimated persistence of reserves and lowers the reserve-adequacy column of the level factor, from 0.041 to 0.022, and the attribution falls with it; under either estimator reserve adequacy accounts for a large share of the variance of the premium. The stationarity adjustment does not bind at the point estimate.

Table E.3 prices reserve adequacy as a spanned factor: the pricing state is extended to four latent factors, reserve adequacy is an exact affine function of them with free coefficients, and the short rate is free to load on it, the normalization of Joslin et al. (2013). The cross-section assigns reserve adequacy a loading of at most 0.15 basis points per standard deviation at any maturity, the twelve-month premia of the two models have a correlation of 0.999, and their loadings on reserve adequacy differ by about half a basis point or less per standard deviation. The spanned model fits the ladder marginally better,

but its likelihood gain is no larger than that obtained by pricing a time-reversed copy of the series or a linear trend in its place, so the gain reflects an additional persistent factor rather than reserve information. The two models are not nested in the regular sense, since the unspanned model is a boundary case of the spanned one, and no p-value is reported.

The largest eigenvalue $\hat{g}_1 = 1.00041$ exceeds unity by about three standard errors, a familiar consequence of estimating a Gaussian model over a sample half of which sits at the zero lower bound, where the level factor has no mean reversion to reveal; the likelihood is nearly flat along a ridge in (g_1, κ), whose estimated correlation is -0.71. We therefore report risk-neutral quantities only within the twelve-month ladder, where g_1^n is numerically indistinguishable from one, and no unconditional quantity at all. On the post-2016 subsample, which drops the first zero-lower-bound episode, the largest eigenvalue falls below one.

The second and third eigenvalues sit at the separation floor and the corresponding state coordinates are collinear (sample correlation -0.9999), so nothing defined coordinate by coordinate within that pair is estimated by the data. Aggregates on the subspace as a whole are invariant, as Joslin et al. (2011) require: re-estimating the risk-neutral block as a diagonal system, a Jordan block, and a complex-conjugate pair leaves the pricing errors, the physical spectrum, the specification statistic, and the term-premium attribution unchanged to within half a percent (Table D.1).

The stripped current-delivery series is the one measurement whose construction touches realized funds-rate fixings, and hence the one channel through which a mechanical link between reserve conditions and the pricing block could enter. Re-estimating without it moves the risk-neutral eigenvalues in the fourth decimal place and κ by about two percent, so no such link operates.

Two curve factors do not suffice to price the ladder: the two-factor model matches the interior but misses both ends by about six and a half basis points, against about two for three factors (Figure 3). The failure is a property of the curve, not of the estimator,

Table D.1: Invariance of the Reported Objects to the Treatment of the Repeated Root

The risk-neutral block is re-estimated under three treatments of the near-repeated second and third eigenvalues: the baseline diagonal parameterization with the separation floor binding, a Jordan block with a genuinely repeated root, and a complex-conjugate pair. Aggregates defined on the two-dimensional subspace are invariant, as they must be under the normalization of Joslin et al. (2011); quantities defined coordinate by coordinate within it are not, and none is reported. Shares are on the least-squares track, as in Table 7. Each column is a full re-estimation of the risk-neutral block under that parameterization, with the physical side and the attribution recomputed from the states it implies. The likelihood prefers the complex pair by about 85 log-likelihood units, and the reported aggregates move by less than half a percent.

	Diagonal (at the floor)	Repeated root (Jordan)	Complex pair
Cross-sectional fit, 1–12 months	unchanged to 3–5 significant figures		
Spectrum of $\mathbf{\Phi}^{\mathbb{P}}$	unchanged to 3–5 significant figures		
Specification test statistic (Section 4)	11.03	11.03	11.00
Reserve-adequacy share of the term premium	0.789	0.789	0.793

since the residual left by two unrestricted principal components is already that large at the ends. In this lower-dimensional model the restriction $\mathbf{\Phi}^{\mathbb{P}}_{X,\mathrm{RA}} = 0$ is rejected more sharply, a likelihood ratio of 12.80 on two degrees of freedom against 11.03 on three. The two-factor rejection is reported as a cross-check rather than as the headline, since a term premium measured in a model that misprices the ends of the ladder by six basis points would not be credible.

E Additional Tables and Figures

This appendix collects the tables and figures referred to in the text: the horizon profile and out-of-sample forecasts under alternative windows (Table E.1, Figure E.1), alternative measures of reserve adequacy (Table E.2), the spanned specification (Table E.3), state dependence of the excess-return loading (Table E.4), reserve adequacy and future repo-market tightness (Table E.5), and the SOFR replication (Table E.6). The last subsection reports the results of Section 2 for the broader measure of reserve adequacy that includes

overnight reverse repurchase balances.

Table E.1: Out-of-Sample Forecasts: Alternative Windows, Subperiods, and Factor Construction

This table repeats the out-of-sample comparison of the curve plus reserve adequacy m against the curve alone (Table 5, block (a)) under alternative estimation windows, forecast periods, curve specifications, and constructions of the curve factors. The target is the futures excess return $rx^{(h)} = f -$ realized in basis points at horizons of one to four quarters; a positive value is compensation to a long position. The curve is the first two principal components of the twelve-contract ladder, and m is the demeaned log ratio of reserves to GDP. "Rolling five-year window" re-estimates each forecasting regression on the five years before the forecast date, with forecasts from January 2015 through June 2025. "Forecasts from 2017" uses an expanding window with the first forecast in January 2017. The two subperiod rows use the expanding-window forecasts of the baseline and evaluate only those made in 2015–2019 or in 2020–2025. "Three-component curve" compares the first three principal components plus m with the first three components, expanding window. The baseline computes the principal-component loadings once on the full sample, as in Joslin et al. (2014); the last two rows remove any use of future prices from the factor construction. "Loadings fixed on 2010–2014" estimates the loadings and the maturity means on the initial training sample, January 2010 through December 2014, and holds them fixed; "Loadings re-estimated at each date" re-estimates them on the data available at each forecast date. In every case the training sample at a forecast date contains only returns already realized by that date. R^2 is the out-of-sample R^2 in percent, $100 \times (1 - \mathrm{SSE}_{\mathrm{model}}/\mathrm{SSE}_{\mathrm{benchmark}})$; CW is the one-sided Clark–West statistic with Newey–West standard errors with $13h$ lags, with one-sided critical values of 1.28, 1.65, and 2.33 at the 10, 5, and 1 percent levels. The numbers of forecasts are 548 at one to three quarters and 539 at four for the rolling window, 443 at one to three quarters and 434 at four for forecasts from 2017, 261 for 2015–2019, 287 at one to three quarters and 278 at four for 2020–2025, and 548 at one to three quarters and 539 at four for the three-component curve and the two real-time-loadings rows.

	1 quarter		2 quarters		3 quarters		4 quarters	
	R^2	CW	R^2	CW	R^2	CW	R^2	CW
Panel A. Estimation window and forecast period								
Rolling five-year window	6.8	2.37	37.4	1.62	46.5	1.52	50.3	1.39
Forecasts from 2017	9.6	2.76	25.7	2.39	32.7	2.12	35.1	1.90
Forecasts 2015–2019	−0.4	0.13	4.1	2.22	9.6	2.84	10.4	2.72
Forecasts 2020–2025	11.4	2.88	33.8	2.51	43.9	2.22	51.2	1.98
Panel B. Curve specification and factor construction								
Three-component curve	7.8	2.36	21.9	2.20	29.1	1.93	31.5	1.65
Loadings fixed on 2010–2014	9.8	2.74	26.7	2.33	34.6	2.09	36.8	1.92
Loadings re-estimated at each date	9.3	2.67	25.6	2.34	33.6	2.12	36.6	1.93

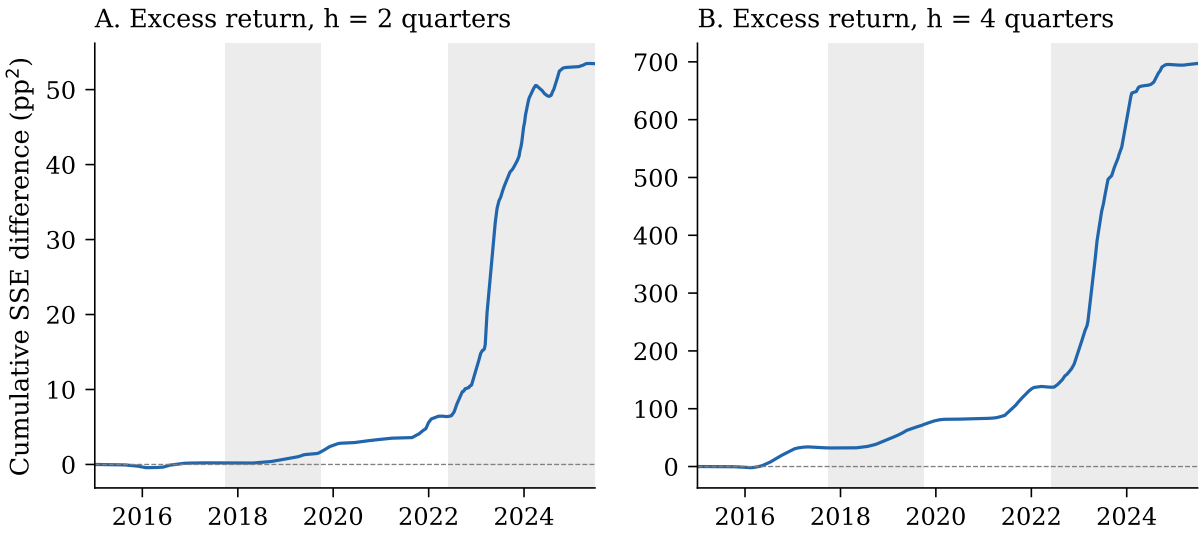


Figure E.1: Cumulative out-of-sample gain from reserve adequacy. Each panel plots the cumulative sum, over forecast dates, of the squared forecast error of the curve-only model minus that of the model that adds reserve adequacy m, in squared percentage points, following Welch and Goyal (2008); a rising line means the model with m is forecasting better. The target is the futures excess return $rx^{(h)} = f -$ realized at two quarters (Panel A) and four quarters (Panel B). The curve is the first two principal components of the twelve-contract ladder. Forecasts are made weekly from January 2015 through June 2025 from regressions estimated on an expanding window that starts in January 2010, using only returns realized by the forecast date. Shaded bands mark balance-sheet runoff, October 2017 to September 2019 and June 2022 to the end of the sample.

E.1 Reserves and the Non-Bank Buffer

Reserve balances are the banking system's claim on the Federal Reserve, but they are not the only overnight claim the private sector holds. Since September 2013 money-market funds, government-sponsored enterprises, and dealers have been able to place cash in the overnight reverse repurchase facility at an administered rate, and those balances are a second outlet for the liquidity the Fed's asset holdings create. The two outlets are substitutes at the margin: when Treasury bill issuance rose in 2023, money funds moved out of the facility and into bills while bank reserves were little changed, and reserves began to fall only once the facility had been largely drained. The quantity that measures how far the system is from reserve scarcity is therefore arguably the sum of the two, reserves plus overnight reverse repurchase balances, rather than reserves alone (Logan, 2023).

Table E.7 reports the results of Section 2 for this broader measure beside the headline measure. The broader measure is unspanned to the same degree, forecasts the level factor of the curve one week ahead with a larger coefficient and a larger t-statistic, carries a larger four-quarter excess-return loading, and forecasts the repo spread over the following quarter with a smaller standard error. Two further rows show how the two measures relate. Adding the level of overnight reverse repurchase balances as a control to the one-step forecast absorbs the coefficient on reserves over GDP, and excluding the 2021–2023 period in which those balances exceeded five percent of GDP leaves the coefficient on reserves over GDP positive and significant at the 5 percent level. Reserves over GDP and the facility's balances are two parts of one liquidity state, and the broader measure captures it with less noise. We keep reserves over GDP as the headline measure because it is the yardstick the policy discussion uses and because the mechanism of Section 5 runs through bank reserves; each result in Table E.7 is at least as strong on the broader measure.

Table E.2: Alternative Measures of Reserve Adequacy

This table repeats three results of the paper with alternative measures of reserve adequacy. Each measure is the demeaned logarithm of the ratio named in the row, on a weekly (Friday) grid from 8 January 2010 to 27 June 2025. Bank assets and deposits are from the H.8 release; because H.8 data for a Wednesday are published after the futures close of the following Friday, each H.8 value enters 16 days after its Wednesday reference date. ON RRP is take-up at the overnight reverse repurchase facility, which is zero before September 2013, so measure (4) equals measure (1) before then. The first column reports the R^2 from projecting the weekly innovation in the measure, computed on its two-week average, on the innovations in the first three principal components of the twelve-contract futures ladder (807 weeks). The second and third columns report the coefficient on the standardized measure in one-week-ahead regressions of the change in the first or second principal component, multiplied by 100, on the two current components and the measure; Newey–West standard errors with four lags (807 weeks). The last two columns report the coefficient on the standardized measure, in basis points per standard deviation, in regressions of the futures excess return $rx^{(h)} = f -$ realized at horizons of two and four quarters on the first two principal components and the measure; a positive return is compensation to a long position, so a negative coefficient means that more reserves go with lower compensation. Newey–West standard errors with $13h$ lags (808 weeks at two quarters, 799 at four). $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Spanning innovation R^2	$\Delta\mathrm{PC}_{1,t+1}$ $\times 100$	$\Delta\mathrm{PC}_{2,t+1}$ $\times 100$	$rx^{(2)}$ (bp)	$rx^{(4)}$ (bp)
(1) Reserves/GDP	0.012	3.12^{***}	1.01^{***}	-17.57^{***}	-52.80^{***}
		(0.93)	(0.38)	(5.48)	(16.90)
(2) Reserves/bank assets	0.014	2.82^{***}	0.91^{**}	-16.49^{***}	-47.12^{***}
		(0.98)	(0.37)	(5.56)	(17.18)
(3) Reserves/deposits	0.014	2.54^{**}	0.74^{*}	-15.40^{***}	-40.31^{**}
		(1.13)	(0.38)	(5.77)	(16.39)
(4) (Reserves + ON RRP)/GDP	0.007	3.86^{***}	1.25^{***}	-20.35^{***}	-63.56^{***}
		(0.96)	(0.44)	(5.38)	(16.52)
(5) Log reserves	0.016	3.37^{***}	1.31^{***}	-18.69^{***}	-60.12^{***}
		(0.78)	(0.43)	(5.70)	(18.90)

Table E.3: Pricing Reserve Adequacy as a Spanned Factor

This table compares the paper's model, in which reserve adequacy m is unspanned (it moves the physical dynamics of the pricing factors but not the cross-section of futures prices), with a model in which m is a spanned fourth pricing factor. The unspanned model has three latent factors; the spanned model has four, with m priced exactly as an affine function of them under the normalization of Joslin et al. (2013), so the short rate may load on m. Both models are estimated by quasi-maximum likelihood on weekly federal funds futures prices from 8 January 2010 to 27 June 2025 (808 weeks), with physical dynamics estimated by least squares. Panel A reports the root mean squared pricing error of each contract, in basis points. Panel B reports the response of the term premium, TP $= f - E^P$, at three to twelve months to a one-standard-deviation increase in m (0.25), in basis points, together with the correlation of the two models' twelve-month premia. It also reports, for the spanned model, the largest absolute response of a fitted contract rate to m across the twelve contracts (attained at contract 1) and the response of the model short rate, both holding the principal components of the curve fixed and both in basis points per standard deviation; in the unspanned model both are zero by construction. Panel C reports twice the gain in the quasi-log-likelihood over the unspanned model for the spanned model and for two placebos that replace m with its time-reversed path or with a linear trend scaled to the standard deviation of m. Each adds six parameters. The models are not nested, so no p-value is reported. The unspanned loadings in Panel B equal the component standard deviations of Table 8, signed for a one-standard-deviation increase in m.

	Unspanned		Spanned
Panel A. Pricing error by contract, root mean squared (bp)			
Contract 1	2.10		2.05
Contract 2	2.22		2.19
Contract 3	2.27		2.23
Contract 4	1.95		1.94
Contract 5	1.50		1.49
Contract 6	1.32		1.31
Contract 7	1.37		1.36
Contract 8	1.43		1.41
Contract 9	1.30		1.30
Contract 10	1.13		1.12
Contract 11	0.95		0.95
Contract 12	1.50		1.48
Mean, contracts 1–12	1.59		1.57
Panel B. Term premium and reserve adequacy (bp per standard deviation of m)			
Loading of 3-month premium on m	−4.57		−4.06
Loading of 6-month premium on m	−14.08		−13.89
Loading of 9-month premium on m	−26.72		−26.52
Loading of 12-month premium on m	−40.62		−40.19
Correlation of 12-month premia		0.999	
Largest curve loading on m, absolute	0		0.15
Short-rate loading on m	0		−0.73
Panel C. Quasi-likelihood gain over the unspanned model			
	Gain		Extra parameters
Spanned, reserve adequacy m	244.24		6
Spanned, time-reversed m	290.94		6
Spanned, linear trend	248.25		6

Table E.4: State Dependence of the Reserve-Adequacy Loading

This table reports regressions of the futures excess return $rx^{(h)} = f -$ realized, in basis points, at horizons of h quarters on the first two principal components of the twelve-contract ladder and on reserve adequacy, with weekly (Friday) observations from 8 January 2010 to 27 June 2025. A positive return is compensation to a long position. Panel A enters reserves/GDP in percent. Panel B adds the shortfall of reserves/GDP below 10 percent, $(10 - \text{reserves/GDP})^+$, so the slope below 10 percent is the sum of the two coefficients. Panel C enters m, the demeaned log ratio of reserves to GDP (one standard deviation is 0.25), its interaction with an indicator for balance-sheet runoff, and the indicator itself; runoff covers October 2017 to September 2019 and June 2022 to the end of the sample. Coefficients on m are in basis points per unit of log reserves/GDP. Newey–West standard errors with $13h$ lags. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	$h = 1$	$h = 2$	$h = 3$	$h = 4$
Panel A. Linear in reserves/GDP				
Reserves/GDP (percent)	-1.84^{**}	-5.91^{***}	-11.84^{***}	-18.86^{***}
	(0.86)	(1.87)	(3.60)	(5.99)
R^2	0.150	0.316	0.410	0.440
Panel B. Kink at 10 percent				
Reserves/GDP (percent)	-0.41	-3.14	-9.19	-17.71^{*}
	(0.40)	(2.30)	(5.69)	(9.57)
$(10 - \text{reserves/GDP})^+$	5.72	11.06	10.59	4.62
	(3.70)	(7.79)	(13.45)	(20.02)
R^2	0.179	0.345	0.421	0.441
Panel C. Balance-sheet runoff				
m outside runoff	-22.53^{*}	-61.11^{**}	-106.68^{**}	-167.94^{**}
	(13.03)	(25.97)	(43.98)	(73.20)
$m\times$ runoff	-25.93	-83.68^{*}	-158.68^{**}	-213.34^{**}
	(19.16)	(48.19)	(71.22)	(99.86)
Runoff indicator	-22.10	-36.44	-21.72	-12.39
	(17.14)	(32.02)	(34.02)	(33.85)
R^2	0.193	0.358	0.450	0.468
Observations	808	808	808	799

Table E.5: Reserve Adequacy and Future Repo-Market Tightness

This table reports predictive regressions of SOFR−IOR, the spread of the secured overnight financing rate over interest on reserves, in basis points, averaged over the daily observations in the 13 weeks (Panel A) or 26 weeks (Panel B) after each weekly (Friday) forecast origin. The regressors are the current week's average spread, the first two principal components of the twelve-contract federal funds futures ladder, and reserve adequacy m, the demeaned log ratio of reserves to GDP, standardized; the table reports the coefficient on m in basis points per standard deviation (0.25 log points), so a negative value means that more reserves predict a lower future spread. Origins run from April 2018, when SOFR begins, to 27 June 2025, and a window enters only with at least $3k$ daily observations for a k-week window. The rows add, in turn, the current three-month Treasury bill rate less IOR; the current IOR less the ON RRP offering rate and an indicator for a technical adjustment of IOR within the window; and the TGA and total reverse repurchase balances in percent of GDP with their 13-week changes. The last row replaces the window average with its 90th percentile. In the 2022–2025 sample the technical-adjustment indicator has no variation and drops out. The full, 2018–2021, and 2022–2025 samples have 378, 196, and 182 origins, equivalent to about 29, 15, and 14 non-overlapping 13-week windows and 14.5, 7.5, and 7.0 non-overlapping 26-week windows. Newey–West standard errors with k lags. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Full	2018–2021	2022–2025
Panel A. Average SOFR − IOR over the next 13 weeks			
Baseline	−4.10***	−0.72	−2.39***
	(0.98)	(1.23)	(0.63)
+ Treasury bill − IOR spread	−4.10***	−0.85	−2.11***
	(0.97)	(1.20)	(0.56)
+ IOR − ON RRP rate and technical-adjustment indicator	−3.40***	−0.81	−2.31***
	(0.97)	(1.23)	(0.74)
+ TGA and RRP levels and 13-week changes	−3.67***	0.66	−1.71***
	(1.09)	(1.02)	(0.60)
Outcome: 90th percentile in the window	−5.40***	−1.25	−6.58***
	(1.19)	(1.62)	(1.02)
Panel B. Average SOFR − IOR over the next 26 weeks			
Baseline	−4.22***	−0.55	−4.15***
	(1.07)	(1.31)	(1.38)
+ Treasury bill − IOR spread	−4.22***	−0.73	−4.02***
	(1.07)	(1.37)	(1.44)
+ IOR − ON RRP rate and technical-adjustment indicator	−3.57***	−0.31	−2.25***
	(0.98)	(1.16)	(0.69)
+ TGA and RRP levels and 13-week changes	−3.13***	1.37	−2.29***
	(1.10)	(1.15)	(0.67)
Outcome: 90th percentile in the window	−5.52***	−1.08	−10.31***
	(1.40)	(1.65)	(2.83)

Table E.6: Reserve-Adequacy Slope in Funds and SOFR Futures on the Common Sample

This table reports the reserve-adequacy slope of the futures excess return, $rx = f -$ realized, in federal funds futures and in CME one-month SOFR futures, estimated on the same weeks. The sample is the 266 weeks from March 20, 2020 to June 27, 2025 in which the first six SOFR contracts and the matching funds contracts all clear the open-interest screen; it therefore excludes the 2017–2019 runoff and September 2019. The SOFR panel is built independently of the funds panel from contract-level data, with its own delivery-month verification and settlement anchor, and the excess return of each contract is measured against its own final settlement. For each market and horizon the slope is the coefficient on reserve adequacy (the two-week moving average of log reserve balances over nominal GDP, demeaned) in a regression of the excess return of the contract nearest to the stated horizon on a constant and that market's own first two curve principal components. Slopes are in basis points per one-log-point *decline* in reserve adequacy, the convention of Table 8; one standard deviation of the state is 0.25 log points. Standard errors are Newey–West with $13h$ lags for an h-quarter horizon. The difference between the two slopes is a small fraction of either standard error at both horizons. The SOFR ladder is not deep enough to support longer horizons.

	Horizon	
	3 months	6 months
Panel A. Federal funds futures		
Slope $\hat{\gamma}$ (bps)	−3.4	51.8
Newey–West standard error	(14.9)	(54.8)
t-statistic	−0.23	0.95
Panel B. One-month SOFR futures		
Slope $\hat{\gamma}$ (bps)	−3.7	56.3
Newey–West standard error	(15.5)	(54.0)
t-statistic	−0.24	1.04
Difference, funds minus SOFR (bps)	0.3	−4.5
Weeks	266	266

Table E.7: Reserves and the Non-Bank Buffer: Reserves/GDP against (Reserves + ON RRP)/GDP

This table compares reserve adequacy measured by reserves/GDP with a measure that adds take-up at the Federal Reserve's overnight reverse repurchase (ON RRP) facility, the cash that money market funds hold outside the banking system. Each measure is the demeaned logarithm of the ratio, on a weekly (Friday) grid from 8 January 2010 to 27 June 2025, and enters standardized. ON RRP take-up is zero before September 2013, so the two measures coincide before then. The first row reports the R^2 from projecting the weekly innovation in the measure, computed on its two-week average, on the innovations in the first three principal components of the twelve-contract futures ladder. The one-step rows report the coefficient on the measure in one-week-ahead regressions of the change in the first principal component, multiplied by 100, on the two current components and the measure, over the full sample (807 weeks), from 27 September 2013 (613 weeks), with ON RRP take-up in percent of GDP as an additional regressor, whose coefficient is reported in the row below, and excluding April 2021 to June 2023 (689 weeks); Newey–West standard errors with four lags. The excess-return row reports the coefficient, in basis points per standard deviation, in a regression of the four-quarter futures excess return $rx^{(4)} = f -$ realized on the first two principal components and the measure, with Newey–West standard errors with 52 lags; a positive return is compensation to a long position. The last row reports the coefficient, in basis points per standard deviation, in a regression of the average SOFR−IOR spread over the next 13 weeks on the current spread, the first two principal components, and the measure, with origins from April 2018 and Newey–West standard errors with 13 lags; in this row only, the second measure adds total reverse repurchase agreements from the H.4.1 release, which include the foreign official pool, rather than ON RRP take-up. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Reserves/GDP	(Reserves + ON RRP) /GDP
Spanning innovation R^2, $K = 3$	0.012	0.007
One-step ΔPC_1, full sample	3.12^{***}	3.86^{***}
	(0.93)	(0.96)
One-step ΔPC_1, from 27 September 2013	4.50^{**}	5.01^{***}
	(1.84)	(1.60)
One-step ΔPC_1 with ON RRP/GDP: measure	1.54	1.76
	(1.23)	(1.46)
ON RRP/GDP	2.72^{**}	2.41^{*}
	(1.14)	(1.33)
One-step ΔPC_1, excluding April 2021–June 2023	2.14^{**}	2.39^{**}
	(1.07)	(1.19)
Excess return $rx^{(4)}$ (bp)	-52.80^{***}	-63.56^{***}
	(16.90)	(16.52)
SOFR − IOR, next 13 weeks (bp)	-4.10^{***}	-3.94^{***}
	(0.98)	(0.71)